

Feasibility, Advantages and Disadvantages of BP Neural Network Applied in TCM Constitution Identifications

Xie Fangrong¹, Zhou Xiaoyun¹, Han Liang², Shi Zhongfeng², Chen Guan hao², Huang Haiquan³, Cheng Qi³, Chen Ziqiang³, Hu Jinyuan⁴, Song Yuhong¹, Xu Shu^{5,*}

¹First People's Hospital Affiliated to Guangzhou Medical University, Guangzhou, China

²Health Departments, Guangdong Pharmaceutical University, Guangzhou, China

³Guangdong Yisheng Information Technology Co., Ltd., Guangzhou, China

⁴Guangdong Shengshijianwang Health Management Co., Ltd., Guangzhou, China

⁵Chinese Academy of Sciences University Shenzhen Hospital (Guangming), Shenzhen, China

Email address:

xiefangrong@126.com (Xie Fangrong), selflearner@126.com (Xu Shu)

*Corresponding author

To cite this article:

Xie Fangrong, Zhou Xiaoyun, Han Liang, Shi Zhongfeng, Chen Guan hao, Huang Haiquan, Cheng Qi, Chen Ziqiang, Hu Jinyuan, Song Yuhong, Xu Shu. Feasibility, Advantages and Disadvantages of BP Neural Network Applied in TCM Constitution Identification. *Asia-Pacific Journal of Computer Science and Technology*. Vol. 1, No. 4, 2019, pp. 34-38.

Received: November 15, 2019; Accepted: February 19, 2020; Published: March 23, 2020

Abstract: Objective This study will analyze the feasibility, advantages and disadvantages of the application of BP neural network in TCM physique identification based on the characteristics of TCM physique identification, so as to obtain a more accurate physique typing, which makes the use of TCM for identifying diseases more widely and convenient. science. Methods The feedforward BP neural network model was used to operate the data to construct a BP neural network model suitable for TCM physique identification. We will carry out a questionnaire survey on community people aged 40 to 70 years old in Longjiang Town, Shunde District, Foshan City, Guangdong Province, collect data models through the TCM physique identification form and make predictions on the population's physique; then match the questionnaire content with the final results Among them, 525 sets of data are used as the training set input model, and the remaining 132 sets of data are used as the test set. After error testing and comparison analysis with the classic prediction model. The results show that the BP neural network method can predict the TCM constitution type of the community based on the questionnaire results of the TCM Constitution Identification Form. Conclusion The application of BP neural network in the classification of TCM constitutions has high reliability, simple operation, low cost, and convenient methods suitable for community promotion.

Keywords: BP Neural Network, TCM Constitution Identification, Feasibility

BP神经网络在中医体质辨识中应用的可行性及优缺点

谢方镛¹, 周晓芸¹, 韩亮², 石忠峰², 陈冠豪², 黄海铨³, 程琦³, 陈自强³, 胡金元⁴, 宋雨鸿¹, 徐舒^{5,*}

¹广州医科大学附属市一人民医院, 广州, 中国

²广东药科大学, 健康学院, 广州, 中国

³广东易生活信息科技有限公司, 广州, 中国

⁴广东盛世健王健康管理有限公司, 广州, 中国

⁵中国科学院大学深圳医院(光明), 深圳, 中国

邮箱

xiefangrong@126.com (谢方镛), selflearner@126.com (徐舒)

摘要: 目的 本研究将根据中医体质辨识特点, 分析研究BP神经网络在中医体质辨识中应用的可行性及优缺点, 从而得出较为准确的体质分型, 使中医辨病的使用更加的广泛, 方便, 科学。方法 运用前馈型的BP神经网络模型对数据进行运算, 构造适用于中医体质辨识的BP神经网络模型。我们将对广东省佛山市顺德区龙江镇年龄40到70岁的社区群众进行问卷填写调查, 通过中医体质辨识表收集数据模并对人群的体质作出预测; 然后将问卷内容及最后结果配对, 其中525组数据作为训练集输入模型, 剩余132组数据作为测试集, 经过误差检验和与经典预测模型进行对比分析。结果 说明BP神经网络方法可以根据《中医体质辨识表》的问卷结果对社区群众的中医体质分型进行预测。结论 BP神经网络在中医体质分型的运用因其结果具有较高的可信度且操作简便、成本低廉、方法便利适用于社区推广。

关键词: BP神经网络, 中医体质辨识, 可行性

1. 中医体质的研究进展

《黄帝内经》根据“阴阳、五行、形态、心理”最早将体质分门别类, 《灵枢·通天》中就曾记载: “盖有太阳之人、少阴之人、太阳之人、少阳之人、阴阳和平之人, 凡五人者, 其态不同, 其筋骨气血各不等”, 将人按阴阳体质不同进行区分[1]。体质的形成受先天、年龄、性别、精神状态、生活及饮食条件、地理环境、疾病、体育锻炼、社会等众多因素的影响。体质与身体对外界刺激的敏感性及其疾病的发生发展都密切相关, 不同体质的人对同一疾病的易感性不同, 发病的症状和程度也不同, 因此治疗时也有不同的方法和禁忌[2-5]。所以体质的分型对疾病的诊断治疗起着关键的作用。

中医体质学是研究人类体质类型的生理、病理特点, 分析病因病机及发展趋向, 指导疾病预防、治疗以及养生康复的一门学科。中医体质辨识是根据不同体质的特征, 将人的健康状况分门别类, 辨别不同证候, 并根据不同体质类别, 制定相应的治疗、预防、养生的方法[6]。目前, 临床上区分人体体质的方法主要仍是依靠有经验的中医师对患者进行望、闻、问、切, 通过患者自身的描述和医生对患者机体状态的判断从而进行体质分类[7,8]。这种方法操作繁琐、过程复杂, 对医生自身素质及与患者的沟通交流方法有着较高要求, 只适用于医院等临床场所, 不易在社区推广。中医体质辨识是治未病的重要方法和手段, 是现代预防医学的基本指导思想[9]。体质分型最重要的作用之一便是“治未病”, 所谓“未病先防、已病防变、已变防渐”, 其中“未病先防”是最基础的部分也是人口基数最大的部分。该部分的人群主体主要集中在社区, 若是可以在此阶段就对人们进行体质分型, 则将能更加有针对性地“防治未病”[10]。

2. BP神经网络

BP神经网络是一种按照误差逆向传播算法训练的多层前馈神经网络, 其网络结构是由输入层、输出层和隐藏层组成的, 由Rumelhart和McCoelland为首的团队在1986年提出[11]。该模型的提出极大地缩小了多层前馈网络模式分类器与实际运用的距离, 并在模式识别、信号

处理、自动控制、辅助决策等多个领域都取得了丰硕的成果[12]。其网络结构是由输入层、输出层和隐藏层组成的, 每一层神经元的作用不同的: 输入层接收外界信息, 输出层对输入层信息进行判别和决策, 中间的隐藏层则用来表示或存储信息[13]。在此, 我们将运用BP神经网络系统的并行处理能力、分布存储能力、容错能力、自适应性等特性, 训练该模型对数据进行运算, 得出较为准确的体质分型。

BP神经网络的工作流程如下: 对每个训练样本, 先将输入特征提供给输入层神经元, 然后逐层将信号前传, 直到产生输出层的结果; 然后计算输出层的误差, 再将误差逆向传播至隐层神经元; 最后根据隐层神经元的误差来对连接权和阈值进行调整, 该迭代过程循环进行, 直到达到某些停止条件为止[14, 15]。

本研究将根据中医体质辨识特点, 运用BP神经网络系统的训练对模型数据进行运算, 构造适用于中医体质辨识的BP神经网络模型, 从而得出较为准确的体质分型, 使中医辨病的使用更加的广泛, 更加方便以及更加科学。

3. BP神经网络预测模型工作步骤

BP神经网络是一种单向传播系统, 它含有多层传播网络, 其输入向量和输出向量的非线性映射关系非常高, 如图1。BP神经网络的输入向量和输出向量分别是归一化处理的结果和量化的学习效果, 具体步骤如下:

- (1) 随机选取样本。从足够多的训练样本库中任意选取输入样本, 对应相应的数值期望输出;
- (2) 计算种群个体输入、输出值;
- (3) 计算输出层神经元的误差函数的偏导数;
- (4) 计算隐含层神经元的误差函数的偏导数;
- (5) 修正连接权值及阈值;
- (6) 计算全局误差;
- (7) 根据结果来决定是否结束算法。

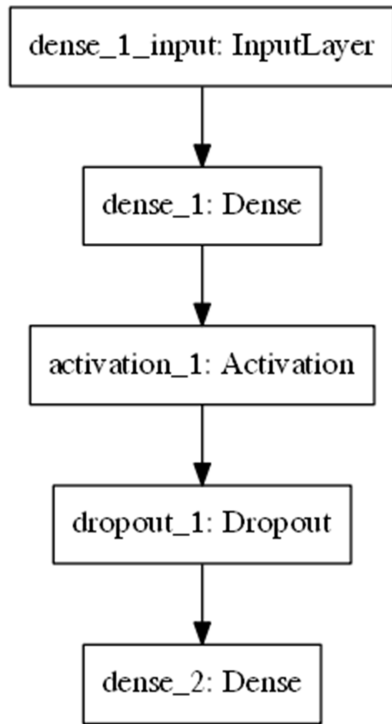


图1 模型概述图。

4. BP神经网络预测模型对社区群众的体质分型方法

4.1. 数据纳入

我们选取广东省佛山市顺德区龙江镇年龄40到70岁的社区群众共657人，其中男性352人，女性305人，分别派发中医体质辨识表。根据其填写的内容，得到每个人阳虚质、阴虚质、气虚质、痰湿质、湿热质、血淤质、特禀质、气郁质、平和质、这9种体质的分数。随后，请多位资深中医专家通过双盲的方法将每种体质分数按北京中医药大学王琦教授的“身心统一论”进行评分和分型，得到该人的体质结果，每位专家之间判断差异无统计学意义。我们将问卷内容及最后结果配对，将其中525组数据作为训练集输入模型，让其进行学习，剩余132组数据作为测试集，对其进行校验。模型因子为阳虚质得分、阴虚质得分、气虚质得分、痰湿质得分、湿热质得分、血淤质得分、特禀质得分、气郁质得分及平和质得分，预测标签为“体质类型”。

4.2. 数据前处理

在此657个人的数据集中，体质类型的多达106种。所以，我们将预测标签从体质类型这一单一标签变为：阳虚、阴虚、气虚、痰湿、湿热、血淤、特禀、气郁、平和、倾向于阳虚、倾向于阴虚、倾向于气虚、倾向于痰湿、倾向于湿热、倾向于血淤、倾向于特禀、倾向于气郁及倾向于平和这18个标签。分别预测该18个标签的值，然后通过后处理合并成体质类型。（已删除原图2）

4.3. 预测结果

当BP神经网络的迭代次数 (epochs)达到了40次左右，验证集正确率开始下降，验证集损失不再降低。说明该模型迭代40次左右，即可训练出最佳效果。训练过程中各项指标的统计图如图2、图3、图4、图5。

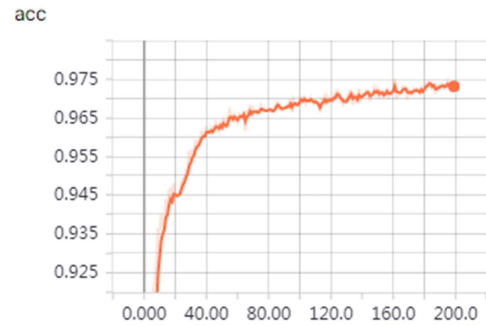


图2 训练集正确率 (trainingaccuracy)。

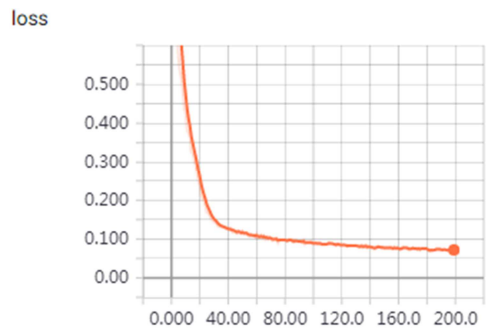


图3 训练集损失 (trainingloss)。

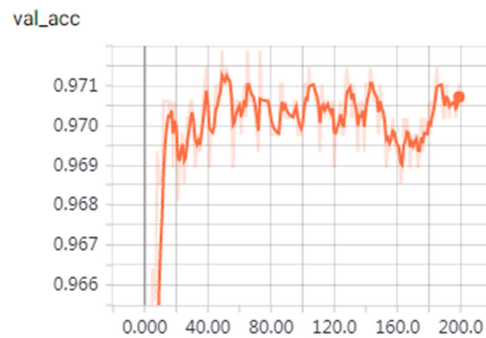


图4 验证集正确率 (validationaccuracy)。

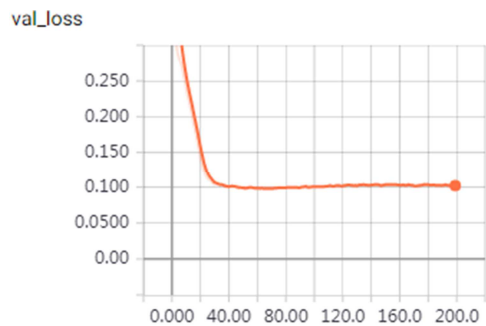


图5 验证集损失 (validationloss)。

为了验证BP神经网络的预测能力,我们将其与经典预测模型主成分分析(Principal Component Analysis, PCA)+逻辑回归(Logistic Regression)及偏最小二乘法判别分

析(Partial least squares discrimination analysis, PLSDA)进行对比。

三种模型用相同的数据进行学习并用相同的数据进行检测及校验,得到结果如下:

表1 相同的数据的三种不同模型之间的检测及校验。

	PCA+Logistic Regression	PLSDA	BP神经网络
Accuracy score	0.765	0.629	0.758
Hamming loss	0.033	0.109	0.055

(1)BP神经网络与PLSDA预测对比

1) Accuracy score得分:神经网络预测准确个数为(132×0.757=100),PLSDA预测准确个数为(132×0.628=83)。

2) Hamming loss得分:神经网络预测错误的标签数是(0.0547×132×18=129),PLSDA预测错误的标签数是(0.109×132×18=259)。

对比BP神经网络和PLSDA:神经网络的准确性和偏差性分数均优于PLSDA。

(2)BP神经网络与PCA+Logistic Regression预测对比

1) Accuracy score得分:神经网络预测准确个数为(132×0.757=100),PCA+Logistic Regression预测准确个数为(132×0.765=101)。

2) Hamming loss得分:神经网络预测错误的标签数是(0.0547×132×18=129),PCA+Logistic Regression预测错误的标签数是(0.033×132×18=78)。

对比神经网络和PCA+Logistic Regression:神经网络的准确性和PCA+Logistic Regression相近,预测的偏差性大于PCA+Logistic Regression。

经过对比,我们发现在预测方面,BP神经网络要优于PLSDA,但在预测的偏差性方面略逊色于PCA+Logistic Regression。所以BP神经网络在中医体质分型的运用因其结果具有较高的可信度且操作简便、成本低廉、方法便利适用于社区推广。

5. 结论

本文运用前馈型的BP神经网络模型对广东省佛山市顺德区龙江镇,年龄40到70岁的社区群众进行了模型学习和体质预测,经过误差检验和与经典预测模型进行对比分析,说明BP神经网络方法可以根据《中医体质辨识表》的问卷结果对社区群众的中医体质分型进行预测。研究的主要结论如下:

- (1) 虽然中医体质辨识起源早并且在疾病诊断和处理上起着重要作用,但因其对医生自身素质要求较高,故而应用范围较为局限。本文根据北京中医药大学王琦教授的“九分法”理论,将BP神经网络模型运用到中医体质辨识中,大大提高了中医体质辨识的效率,也使得中医体质辨识在社区更易推广,社区群众“治未病”更易实现。
- (2) 为了验证BP神经网络的可行性和可靠性,本文除了令BP神经网络模型学习数据并用数据校正及检验外,还用经典的预测模型PLSDA及PCA+Logistic Regression做同步学习和预测与之对比。我们发现在

数据预测的准确性方面,BP神经网络优于其他两种预测模型。在利用实例与标记间的关系方面,优于PLSDA,稍逊于PCA+Logistic Regression但在可接受范围内。并且由于其比PCA+Logistic Regression更方便、低廉,故BP神经网络作为中医体质分型的预测模型在社区的运用有更大的前景。

基金项目

广州市高校创新创业教育项目(201709T15),广州市高校创新强校工程项目(Q17024006),深圳市光明区中医药科研项目(GM2019020001)。

参考文献

- [1] 朱燕波,严辉,李彦妮,等.中医体质四个基本原理的实证研究概述[J].中医杂志,2018,59(17):1446-1449.
- [2] 陈捷.不同体质类型亚临床甲减患者与动脉硬化的关系研究[J].河北医学,2018,24(09):1475-1478.
- [3] 李华萍,吴冠虹,郑翠红,等.不同体质指数对高热患者药物降温效果的影响[J].上海护理,2018,18(08):55-57.
- [4] 李敏莹,孙伟鹏,王琳.慢性荨麻疹患者不同体质对睡眠指数的影响调查[J].世界最新医学信息文摘,2018,18(76):33-35.
- [5] 吴娜娜.不同体质与慢性代谢性疾病相关性研究进展[J].中医药临床杂志,2018,30(09):1771-1775.
- [6] 王琦,盛增秀.中医体质学说[M].江苏科技出版社,1982:15.
- [7] 俞若熙,王琦,王济,等.体质辨识应用研究现状分析[J].中国中医药信息杂志,2013,20(02):107-109.
- [8] 赵灿,赵彦青,王松龄.王松龄从少阳厥阴体质辨治偏头痛经验浅析[J].中国民族民间医药,2018,27(13):55-57.
- [9] 杨黛仙.体质辨识在健康体检中应用的初步探讨【J】.中医临床研究,2011:101—104.
- [10] 辛彤.浅析中医体质辨识对临床治疗的作用[J].世界最新医学信息文摘,2018,18(64):154-155.
- [11] 王旭,王宏,王文辉.人工神经网络原理与应用【M】.东北大学出版社,2008:51~59.

- [12] 王智平, 刘在德, 高成秀等, 遗传算法在BP网络权值学习中的应用[J]. 甘肃工业大学学报, 2001, 27(2)。
- [13] 袁铸, 申一歌. 农业机器人轨迹优化自动控制研究——基于BP神经网络与计算力矩[J]. 农机化研究,2017(06):33-37。
- [14] 万昊, 谭宗颖, 张福俊, 等. 项目验收的同行评议辅助决策评价方法研究——基于贝叶斯正则化修正的BP神经网络模型[J]. 情报杂志,2017,11(36):1002-1965。
- [15] 王军, 谭继文, 战卫侠. 基于BP神经网络的钢丝绳断丝信号处理[J]. 煤矿机械,2011,32(08):256-258。

作者简介



谢方榕(1993-), 女, 硕士研究生, 研究方向为消化内科中西医结合方向。



徐舒(1975-), 男, 博士, 主任医师, 主要从事中西医结合消化疾病治疗与预防研究。