

复合性国际诊断交谈表计算机神经网络诊断系统的应用研究

邹义壮 舒良 沈渔邨 王玉凤 冯锋 许克勤 屈英 宋燕明

【摘要】 目的 将计算机人工神经网络技术与世界卫生组织(WHO)推出的复合性国际诊断交谈表(CIDI)结合,开发 CIDI 计算机神经网络诊断系统(CIDI-ANN),以辅助精神疾病的诊断。方法 建立误差反向传播网络(BP 网络)和自组织映射网络(kohonen 网络)辅助 CIDI 诊断神经症和精神分裂症。使用 60 例样本训练神经网络,建立 CIDI-ANN,使用另外的 222 例样本进行跨效度检验。结果 对 WHO 提供的 CIDI 诊断软件 CIDI/国际疾病分类第 10 版、CIDI/精神障碍诊断与统计手册第 3 版修订本与新建立的 CIDI-ANN 进行平行测试,3 个系统与临床诊断的总符合率的 Kappa 值分别为 0.77, 0.79 和 0.94,对疑难病例的诊断符合率的 Kappa 值分别为 0.48,0.33 和 0.69。结论 在辅助 CIDI 诊断中,人工神经网络方法优于传统的人工智能方法。人工神经网络与诊断量表的结合是一个新的有前途的研究方向。

【关键词】 精神病学 诊断,计算机辅助

Composite International Diagnostic Interview artificial neural network diagnostic system ZOU Yizhuang*, SHU Liang, SHEN Yucun, et al, *Beijing Huilongguan Hospital, Beijing 100096

【Abstract】 Objective Artificial Neural Network(ANN), as a potential powerful classifier, was explored to assist psychiatric diagnosis of the Composite International Diagnostic Interview(CIDI). **Methods** Both Back-Propagation(BP) and Kohonen networks were developed to fit psychiatric diagnosis and programme (using 60 cases) to classify neurosis, schizophrenia and normal people. The programmed networks were cross-tested using another 222 cases. All subjects were randomly selected from two psychiatric hospitals in Beijing. **Results** Compared to ICD-10 diagnosis by psychiatrists, the overall kappa of CIDI/ICD-10, CIDI/DSM-III-R and CIDI-ANN were 0.77, 0.79 and 0.94. In classifying patients who were difficult to diagnose, the kappa were 0.48, 0.33 and 0.69 respectively. **Conclusions** ANN was more powerful than traditional expert system and might be a new method to improve psychiatric diagnosis.

【Key words】 Psychiatry Diagnosis, computer-assistant

人工神经网络(artificial neural network, ANN)是 80 年代后期迅速发展起来的一门新兴学科,其核心是使用计算机来模拟人脑神经网络的结构与功能,以达到提高计算机智能的目的。人工神经网络是由大量的非线性计算单元按照生物神经网络的结构并行排列,分层连接,能够完成许多“智能”型的工作。精神疾病的诊断过程需要对大量多维的非线性和不精确信息进行处理并作出决策,其工作具有很强的经验性、模拟性和模糊性;传统的统计学和专家系统方法在处理这类信息时的敏感性、准确性和可操作性都较差,而人工神经网络作为在分类、预测和辅助

决策领域的新方法,擅长于处理此类的问题^[1]。

世界卫生组织和美国酒精药物滥用和精神卫生署合作编制了复合性国际诊断交谈表(Composite International Diagnostic Interview, CIDI),并发表了能够分别根据国际疾病分类第 10 版(ICD-10)和美国精神疾病诊断与统计手册第 3 版修订本(DSM-III-R)进行诊断的计算机软件(CIDI/ICD-10 和 CIDI/DSM-III-R)^[2,3]。国内现场测试提示,CIDI 的诊断敏感性偏低(67%~88%)^[4],原因之一是 CIDI/ICD-10 程序使用传统人工智能的逻辑决策树的方法,在模拟医生的诊断经验和处理模糊性较大的 CIDI 数据时有较大的局限性。本研究的目的是将最新的人工神经网络技术引入精神疾病的诊断领域,建立 CIDI 计算机神经网络诊断系统(CIDI-ANN),以提高 CIDI 的诊断效度。

本研究获北京市科技成果三等奖

作者单位:100096 北京回龙观医院(邹义壮、冯锋、许克勤、屈英、宋燕明);北京医科大学精神卫生研究所(沈渔邨、舒良、王玉凤)

资料和方法

一、资料

首先按规定程序对参加研究的 5 名精神科医生进行培训,并进行了测查者间的一致性测验,测查时参加培训的医生轮流担任交谈者和观察者。

1. 患者样本:共 309 例,其中男 166 例,女 143 例,均为随机抽取的北京回龙观医院和北京医科大学精神卫生研究所的住院和门诊患者。要求 2 名精神科医生(其中至少有 1 名主治医师以上的医生)按照 ICD-10 标准独立做出诊断,诊断一致方能入组。309 例中神经症 100 例,男 62 例,平均年龄(30 ± 10)岁,平均受教育(12 ± 3)年;女 38 例,平均年龄(35 ± 14)岁,平均受教育(12 ± 3)年;其中强迫症 51 例、抑郁性神经症 21 例、焦虑症和恐怖症 21 例、神经衰弱 6 例和疑病症 1 例;住院患者占 74%,门诊患者占 26%。精神分裂症 114 例,男 57 例,平均年龄(34 ± 10)岁,平均受教育(11 ± 3)年;女 57 例,平均年龄(33 ± 10)岁,平均受教育(12 ± 3)年;住院患者占 97%。精神分裂症的病程要求达半年以上。

2. 正常样本:共 95 名,主要随机抽取自北京回龙观医院的健康职工。男 47 名,平均年龄(35 ± 11)岁,平均受教育(12 ± 3)年;女 48 名,平均年龄(34 ± 9)岁,平均受教育(13 ± 3)年。

为检验 CIDI 对疑难病例的诊断能力,除以上样本外,特意采集 16 例临床上诊断困难和不典型的病例,其中神经症 10 例,精神分裂症 6 例。疑难病例的标准是:(1)在治疗过程中曾经更改过诊断;(2)症状表现不典型,经科主任或院长查房最后确诊的病例。

全部样本的智力水平均在正常范围(由 CIDI 的 M 节测试)。

二、方法

我们与北京邮电大学钟义信教授及王明会、刘伟权两位博士合作,建立了适用于精神病学使用的计算机人工神经网络系统,包括误差反向传播的多层前馈型神经网络(简称 BP 网络)和 Kohonen 的自组织映射型神经网络(简称 Kohonen 网络),程序使用 C++ 语言编制。为适应精神疾病诊断的需要,在网络规模、训练方式、初始化权值、隐层设置和特征加权等方面做了较大改进,进行了上百次的调整和实验,保证了神经网络结构的最优化,建立了整套的实验室方法。全部试验结果都经过了重复测试,

以保证实验结果的稳定和具有良好的可重复性。

全部数据的统计处理均在计算机上完成,使用的统计软件包为 SPSS/PC+。

结 果

一、检查者间的一致性

参加本项研究的 5 名检查者轮流担任检查者和观察者共同对患者进行检查,总共做了 20 份 CIDI,经统计学处理,检查者间项目评分的一致性为 0.935。

二、CIDI 神经网络系统的诊断效度

使用 CIDI 中的 B~X 节的全部 396 个项目回答作为神经网络的输入,3 个输出端对应正常、神经症和精神分裂症 3 类诊断。对数据首先使用经验模拟的模糊加权方法进行预处理:症状缺失为 1,症状无临床意义及症状可能存在、不清楚或拒绝回答为 2,物质滥用为 3,躯体问题为 4,精神症状为 9。数据的组织是 1 名正常人、1 例神经症和 1 例精神分裂症组成为 1 组,样本随机分为训练集的 20 组 60 例样本和跨效度(cross validity)检验集 74 组 222 例样本。试验中分别使用了 BP 网络和 Kohonen 网络,以比较不同人工神经网络模型的优劣。

1. BP 网络的诊断效度:BP 网络是“有教师指导”的神经网络,研究中使用有丰富临床经验的医生(教师)的诊断结果训练神经网络。训练中网络的输出与医生的诊断进行比较,差异被反向传播回去,对神经网络各节点的权值进行调节,该过程反复自动进行,最终使这一差异成为最小。BP 网络的隐层神经元数目经过优化实验确定为 40 个,各节点的原始权值由计算机随机产生($-0.3 \sim 0.3$),网络训练的学习率为 0.15,训练系数为 1,最大错误水平逐级设置为 0.15,0.1 和 0.05。研究中使用了“成组训练”模式,即每次用 3 个类别同时对网络进行训练,比“单一训练”模式更加有效,结果更加准确。神经网络通过训练可以学会识别正常、神经症和精神分裂症。然后,使用跨效度方法比较 CIDI-ANN 诊断与临床诊断的一致性,CIDI-ANN 系统与临床诊断的跨效度比较情况见表 1。

使用 CIDI-ANN 对 16 例临床医生诊断困难和症状不典型的病例进行了测试,与最终临床诊断的符合率为 81.25%, $Kappa = 0.69, P < 0.01$ 。

2. Kohonen 网络的诊断效度:Kohonen 网络与 BP 网络不同,是两层的自组织映射神经网络,能够在不

知道医生诊断的情况下自动对样本进行组织和分类。为了探讨 Kohonen 网络训练中学习率对诊断结果的影响,试验中将学习率分别设置为 0.6, 0.7, 0.8 和 0.9, 增量为 0.0, 对网络进行优化训练, 每种学习率的训练次数均为 240 次。结果显示, 学习率为 0.6 时的 Kohonen 网络跨效度诊断与临床诊断的符合率最高, 见表 2。

3. CIDI/ICD-10、CIDI/DSM-III-R 和 CIDI-ANN 对神经症、精神分裂症及疑难病例诊断的效度指标比

表 1 CIDI-ANN 诊断与临床诊断比较(例数)

临床诊断	CIDI-ANN 诊断		
	正常	神经症	精神分裂症
正常	73	1	0
神经症	7	67	0
精神分裂症	2	0	72

注: 临床诊断与 CIDI-ANN 的诊断总符合率 = 95.50%, Kappa = 0.94, P < 0.01

表 2 Kohonen 网络诊断与临床诊断比较(例数)

临床诊断	Kohonen 网络诊断		
	正常	神经症	精神分裂症
正常	74	0	0
神经症	18	56	0
精神分裂症	3	1	70

注: Kohonen 网络诊断与临床诊断比较, 总符合率 = 90.09%, Kappa = 0.88, P < 0.01

表 3 不同诊断软件诊断效度的比较

项目	CIDI/ICD-10	CIDI/DSM-III-R	CIDI-ANN
神经症			
例数	100	100	74
敏感性(%)	67.00	74.00	90.54
特异性(%)	99.52	98.56	99.34
符合率(%)	89.00	90.61	96.40
Kappa 值	0.73	0.77	0.92
精神分裂症			
例数	114	114	74
敏感性(%)	88.59	85.96	97.30
特异性(%)	100.00	100.00	100.00
符合率(%)	95.79	94.82	99.10
Kappa 值	0.91	0.89	0.98
区分神经症、精神分裂症和正常人			
例数	309	309	222
符合率(%)	84.79	86.41	95.50
Kappa 值	0.77	0.79	0.94
疑难病例			
例数	16	16	16
符合率(%)	68.75	50.00	81.25
Kappa 值	0.48	0.33	0.69

较: 由于 BP 网络的诊断效度高于 Kohonen 网络, 最终建立的 CIDI 计算机神经网络诊断系统(CIDI-ANN)中使用了 BP 网络模型。研究中使用 CIDI/ICD-10 和 CIDI/DSM-III-R 诊断软件对 309 例 CIDI 样本进行了诊断, CIDI/ICD-10、CIDI/DSM-III-R 和 CIDI-ANN 对神经症、精神分裂症及疑难病例诊断的效度指标比较见表 3。

讨 论

Janca 等^[5,6]报道了 20 例患者使用 CIDI 诊断的结果: CIDI/DSM-III-R 诊断与临床诊断的总符合率 Kappa 值为 0.78, 其中抑郁症的符合率 Kappa 值为 0.84, 焦虑症和恐怖症为 0.76; CIDI/ICD-10 诊断与临床诊断的总符合率 Kappa 值为 0.77, 其中抑郁症的 Kappa 值为 0.78, 焦虑症和恐怖症为 0.73。邹义壮等^[4]在国内对 309 例样本使用 CIDI/ICD-10 和 CIDI/DSM-III-R 进行测试, 与临床诊断的总符合率的 Kappa 值分别为 0.77 和 0.79。舒良等使用 CIDI/ICD-10, 对精神分裂症和情感性精神疾病进行了现场测试, 与临床诊断的符合率分别为 84.2% 和 80.2%^[7]。通过比较可以看出, CIDI 在中国和美国具有近似的诊断效度, 说明不同民族和文化的差异对 CIDI 的诊断效度影响不大。我们使用人工神经网络方法辅助 CIDI 诊断, 使 CIDI 的诊断效度明显提高(Kappa = 0.94), 提示计算机诊断方法的改进可以提高 CIDI 的诊断效度。

人工神经网络与 CIDI/ICD-10 和 CIDI/DSM-III-R 中使用的专家系统逻辑决策树方法相比具有以下优越性: (1) 前者使用了多维推理, 而后者只是一维推理, 因此, 神经网络更适用于模拟和解决复杂的精神疾病诊断问题。(2) 前者可以对医生的诊断经验进行比较理想的模拟, 除逻辑思维外, 还包括专家本人无法用语言表述但在诊断中起重要作用的形象、直觉和经验思维的成份。而后者则只能模拟有限的逻辑思维, 对其它思维形式则无能为力。(3) 前者具有擅长处理模糊、不完全和不精确数据的能力, 因此, 特别适合于处理精神疾病诊断量表中的数据, 而后者处理这类数据的能力较差^[1]。

研究中对 CIDI 的 3 种计算机诊断程序 CIDI/ICD-10、CIDI/DSM-III-R 和 CIDI-ANN 进行了测试, 以两名有经验的临床医生的诊断作为金标准, 经过跨效度的平行测试和 Kappa 检验, 比较了不同系统的诊断效度。结果显示: CIDI-ANN 的诊断效度优于

CIDI/ICD-10 和 CIDI/DSM-III-R, 对临床疑难病例诊断时, 人工神经网络的优越性更加明显(表 3)。

本结果还显示, 人工神经网络中 BP 网络的诊断效果(符合率 $Kappa = 0.94$) 优于 Kohonen 网络(符合率 $Kappa = 0.88$)。这与迄今为止对人工神经网络不同算法模型比较研究结果一致^[1]。

应当指出, 人工神经网络本身的“黑箱”特征限制了系统对诊断结果的解释功能, 这有待于今后对神经网络的理论研究来解决。此外, 我们只对神经症、精神分裂症和正常人进行了区分, 今后的研究中应当增加诊断的类别, 以提高系统的实用价值。我们认为, 人工神经网络可以提高 CIDI 的诊断效率, 优于传统的人工智能方法。CIDI-ANN 对于临床常见的神经症和精神分裂症具有令人满意的诊断效率, 在临床和研究中有一定的实用价值。人工神经网络技术可能成为精神病学研究中的一种新方法。

参 考 文 献

- 1 钟义信, 潘新安, 杨义先. 智能理论与技术——人工智能与神经网络. 北京: 邮电出版社, 1993. 406-488.
- 2 WHO. 复合性国际诊断交谈检查(CIDI)核心本检查者用本、使用者手册和计算机使用手册. 北京医科大学精神卫生研究所译. 1991.
- 3 Robins LN, Wing J, Wittchen HU, et al. The composite international Diagnostic Interview. Arch Gen Psychiatry, 1988, 45: 1069-1074.
- 4 邹义壮, 沈渔邨, 舒良, 等. 精神疾病诊断的方法学研究——量表诊断与人工神经网络[学位论文]. 北京: 北京医科大学精神卫生研究所, 1994.
- 5 Janca A, Robins L N, Buchholz KK, et al. Comparison of Composite International Diagnostic Interview and Clinical DSM-III-R criteria Checklist Diagnosis. Acta Psychiatr Scand, 1992, 85: 440-443.
- 6 Janca A, Robins LN, Cottler LB, et al. Clinical observation of assessment using the Composite International Diagnostic Interview, CIDI analysis of the CIDI Field Trials - Wave II at the St Louis site. Br J Psychiatry, 1992, 160: 815-888.
- 7 舒良, 张鸿燕, 汪向东, 等. “复合性国际诊断交谈检查—核心本”现场测试——信效度检验. 中国心理卫生杂志, 1993, 7: 208-210.

(收稿: 1998-01-12 修回: 1999-01-28)

(本文编辑: 杨小昕)

· 论 著 摘 要 ·

苗、汉两族人群对饮酒认识的差异

金泰安 万学东 欧阳杏娟 袁平 曹玉鸣

为探讨苗、汉两族在饮酒认识上的差异, 我们用自行设计的饮酒认识问卷(以下简称问卷)于 1992 年 4 月至 1993 年 4 月在贵州省农村进行抽样调查。

1. 对象和方法: 对象为贵阳、黔南、黔东南、安顺、遵义等地农村苗、汉两族(相邻但分别集居) ≥ 20 岁的人群。按随机整群抽样方法抽取 10 个调查点, 每点抽查两族男女各 50 人, 两族各调查 1 000 人。两族的年龄、婚姻、经济状况差异无显著性($P > 0.05$)。

2. 调查工具: 由 16 题组成的饮酒认识问卷, 包括饮酒行为认识、饮酒功能认识、饮酒量认识 3 个因子, 以“是”或“否”回答。其中对饮酒持否定态度称负向问题, 对饮酒持肯定态度称正向问题。

调查者均系主治医师以上职称的精神科医生, 并经一致性培训。由调查者逐人逐题地根据被调查者的回答判定

并记录。数据采用 χ^2 检验。

3. 结果: 16 题回答为“是”的题中, 有 11 题苗、汉两族之间的差异有显著性, 其中认为以一醉方休为快, 苗族有 188 人, 汉族有 153 人; 认为饮酒可解乏, 苗族有 844 人, 汉族有 789 人($\chi^2 = 4.331$ 和 $10.095, P < 0.05$); 还有 9 题两族间的差异呈非常显著性($\chi^2 = 30.350 \sim 135.672, P < 0.01$): 苗、汉两族(以下数字按此顺序)认为饮酒可增进友谊分别为 728 人和 581 人; 可解愁为 534 人和 411 人; 可体现男子气质为 427 人和 285 人; 常误事为 629 人和 768 人; 浪费钱为 625 人和 813 人; 无好处为 627 人和 794 人; 是享受为 564 人和 426 人; 酒量越大越好为 322 人和 108 人; 应控制饮酒量为 768 人和 898 人。

在上述 11 个问题中, 4 个负向问题汉族人数多于苗族, 7 个正向问题苗族人数则多于汉族。表明苗、汉两族在对饮酒的认识上差异存在显著性。在上述

3 个因子所含问题个数中, 苗、汉两族间差异有显著性者与差异无显著性者之比分别为 6:1, 2:4, 3:0, 表明两族间在饮酒行为认识和饮酒量认识上的差异更明显。

讨论: 本次调查使用的问卷在对饮酒的认识上分为正向和负向问题, 主要反映苗、汉两族人群对饮酒所持的态度, 问卷中的 3 个因子便于分析被调查人群间的问题所在。结果显示对饮酒持正面肯定的态度苗族人群多于汉族人群。两族间的差异主要表现在饮酒行为认识和饮酒量认识 2 个因子, 形成了苗、汉两族在饮酒问题上不同的社会文化背景。因此, 在苗族地区进行有关饮酒问题的宣传教育工作时应充分注意苗、汉两族间的差异及其所形成的文化背景。

(收稿: 1998-02-09 修回: 1998-05-11)

(本文编辑: 刘宏林)

作者单位: 550004 贵阳, 贵州省安宁医院