

暴力风险评估中的统计预测方法及其应用

刘元元^{1,2}, 胡峻梅^{2,3}, 杨珉^{1,4}, 李晓松¹

(1. 四川大学华西公共卫生学院卫生统计学教研室, 四川 成都 610041; 2. 中国政法大学 证据科学教育部重点实验室, 北京 100088; 3. 四川大学华西基础医学与法医学院法医精神病学教研室, 四川 成都 610041; 4. 英国诺丁汉大学社区健康科学学院精神科, 英国 诺丁汉 NG7 2TU)

摘要: 如何改进暴力风险评估是一个亟待解决的全球性难题, 通过统计学方法进行犯罪的暴力风险评估是其中不可或缺的一环, 作用十分重要。本文主要从统计学角度对暴力风险评估预测方法进行了回顾, 并对其中较为常用的多元统计模型代表的 Logistic 回归模型、数据挖掘技术代表的决策树模型以及人工智能技术代表的神经网络模型及其应用进行综述, 以期对暴力风险评估的进一步研究提供参考。

关键词: 统计学; 暴力; 综述[文献类型]; 危险性评估

中图分类号: DF795.3 文献标志码: A doi: 10.3969/j.issn.1004-5619.2013.03.016

文章编号: 1004-5619(2013)03-0216-06

Statistical Prediction Methods in Violence Risk Assessment and Its Application

LIU Yuan-yuan^{1,2}, HU Jun-mei^{2,3}, YANG Min^{1,4}, LI Xiao-song¹

(1. Department of Health Statistics, School of Public Health, Sichuan University, Chengdu 610041, China; 2. Key Laboratory of Evidence Science, China University of Political Science and Law, Ministry of Education, Beijing 100088, China; 3. Department of Forensic Psychiatry, School of Basic and Forensic Medicine, Sichuan University, Chengdu 610041, China; 4. Division of Psychiatry, School of Community Health Sciences, the University of Nottingham, Nottingham NG7 2TU, UK)

Abstract: It is an urgent global problem how to improve the violence risk assessment. As a necessary part of risk assessment, statistical methods have remarkable impacts and effects. In this study, the predicted methods in violence risk assessment from the point of statistics are reviewed. The application of Logistic regression as the sample of multivariate statistical model, decision tree model as the sample of data mining technique, and neural networks model as the sample of artificial intelligence technology are all reviewed. This study provides data in order to contribute the further research of violence risk assessment.

Key words: statistics; violence; review [publication type]; risk assessment

暴力及暴力犯罪是一个全球性的公共卫生问题, 无论任何法律体系和政府, 均面临对暴力及暴力犯罪行为的预防与管理, 涉及暴力风险评估。暴力风险评估是通过量化的方法对暴力事件发生及其带来的影响和损失的可能程度进行预测和评估。其预测评估效果好坏将直接影响到社会稳定及民众安全, 具有重要意义。近三十年在暴力风险评估的研究中, 较多集中于对预测因子的开发和验证^[1], 相对而言, 对统计方法的研究较少。众所周知, 除客观准确的临床判断外, 所

采用的统计预测方法也将在很大程度上决定评估结果的准确性。因此, 本文主要从统计学角度对暴力风险评估的预测方法进行回顾, 对常用的 Logistic 回归、决策树和神经网络等多元统计模型和数据挖掘技术及其应用进行综述, 以期对今后暴力风险评估的深入研究提供参考。

目前在暴力及暴力犯罪预测研究中采用的统计方法较多, 如受试者工作特征(receiver operating characteristic, ROC) 曲线分析^[2]、传统的 Nuffield 法^[3]、交互分类表(cross-classification tables)、多重回归分析(multiple regression methods)^[4]、多重判别分析(multiple discriminant analysis)、多维列联表分析(multidimensional contingency table analysis)^[5]、Logistic 回归分析、判别分析、聚类分析, 还有如决策树模型(decision tree model)、神经网络(neural networks, NN) 模型等数据挖掘技术等^[6-8]。其中作为传统多元统计方法的经典代表, Logistic 回归模型是目前应用最多且最

基金项目: 国家自然科学基金资助项目(81072513); 证据科学教育部重点实验室开放基金资助项目(2011KFKT02)

作者简介: 刘元元(1979—), 女, 四川乐山人, 讲师, 博士, 主要从事风险评估与决策、分类预测及时间序列预测等研究; E-mail: y_multi@126.com

通信作者: 李晓松, 男, 教授, 主要从事统计方法与卫生决策、统计方法在传染病流行病学、药物临床试验的应用研究; E-mail: lixiaosong1101@126.com

为广泛的方法之一。而近年新兴的数据挖掘技术中,决策树模型以及作为人工智能技术代表的神经网络模型也引起了越来越多的关注。

1 Logistic 回归模型

Logistic 回归模型是对二分类因变量做回归分析时应用最为普遍的多元统计分析方法^[9]。与线性回归模型不同,Logistic 回归模型既不假设因变量和自变量间存在线性关系,也不要求变量服从正态分布或具备方差齐性,自变量可以是连续性变量也可以是分类变量。因此,该方法具有限制条件少、适用范围广的优点^[10]。除了探讨影响因素分析外,Logistic 回归模型还可用于分类预测。其基本原理是将因变量转变成为 logit 变量(即因变量结局是否发生的概率的自然对数),并应用极大似然估计法来估计模型参数^[11],这样 Logistic 回归模型可用于估计某一特定事件(如罪犯释放后的暴力再犯)的发生概率。目前,该方法已被司法精神病学和心理学领域广为认可,用于暴力和其他犯罪行为的预测,并获得了较好的效果^[12-14]。国内王小平等^[15]曾采用 Logistic 回归建模对 150 例精神分裂症患者的攻击行为进行预测,及其相关因素进行了研究,主要包括人格特征、智力结构、脑电图、社会支持和早年不良家庭环境等,并通过结果制定了一个有 21 个条目的攻击行为预测表,其预测准确性达 90.3%,有一定预测效度。

2 决策树模型

决策树模型又称分类树模型(classification tree model),这类方法近年来得到较多研究学者和临床医生的推荐,因为该模型被认为反映了实际临床思考过程,是临床实践中风险判断和决策制定的良好代表^[16]。从最早的自动交互检测(automatic interaction detection, AID)^[17]模型到之后被广为使用的分类回归树(classification and regression tree, CART)模型^[18],均是一种“问题-决定式树模型”(question-decision-tree model)。其基本原理为逐步回答一系列有层次结构性的问题(分割变量),每一步决策紧扣上一问题的答案,直至最终归类(终结点)的获得。其结构示意图呈树状(图 1),故而得名“树”模型。除 CART 模型外,决策树模型还包括卡方自动交互检测(chi-square automatic interaction detection, CHAID)^[19]模型、C4.5 算法(之前的版本是 ID3 算法,之后的版本是 C5.0/See5 算法)^[20-21]、QUEST 算法(quick, unbiased, efficient statistical trees)^[22]、提举树(boosted trees)^[23]模型和迭代分类树(iterative classification trees, ICT)^[24]模型等。

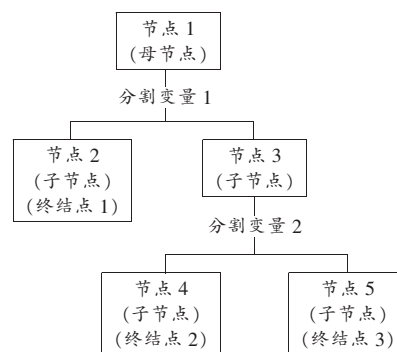


图 1 决策树模型结构示意图

Gardner 等^[12]最先将 CART 模型应用到暴力预测中,他们在一个精神病急诊科患者样本中对负二项回归、一般传统回归模型与 CART 模型进行了比较,并采用了自举法(bootstrapping)进行验证,结果显示 3 种方法的准确性无差异,但他们认为 CART 模型在实践中使用起来更简便。Stalans 等^[25]采用了基于最优判别分析构建的分类树模型和 Logistic 回归模型对 1 344 名暴力犯罪缓刑犯进行了累犯预测,结果发现分类树模型的灵敏度明显高于 Logistic 回归模型,但受试者工作特征曲线下面积(AUC)则显示 Logistic 回归模型略优。Rosenfeld 等^[26]在 204 名犯人对 CART 模型和 Logistic 回归模型进行了比较,但在刀切法(jackknife method)交叉验证后,CART 模型的预测准确性明显降低,而 Logistic 回归模型则表现出了良好的稳定性。类似地,Thomas 等^[13]在 UK700 研究中比较了 CART 和 Logistic 回归模型对精神病患者的暴力预测效果。结果显示在十折交叉验证后,CART 模型的预测表现较 Logistic 回归模型有更多准确性的降低和模型的不稳定。

此外,决策树模型在暴力犯罪预测中的另一著名范例应用是 Steadman 等^[24]自 1998 年起进行的 MacArthur 暴力风险评估研究(MacArthur Violence Risk Assessment Study)中的 ICT 模型。该项研究的主要目的是创建一个可以在精神病患者人群中使用的暴力风险预测和管理临床评估工具。这项前瞻性的多中心项目所包含的样本含量在当时的暴力风险评估领域中最庞大的^[16],研究人员在多个社区内调查随访了 1 100 名精神病患者的暴力犯罪情况。他们希望在该项目中,能克服先前暴力风险评估领域中存在的一些方法学问题,如预测因子的纳入受限制、评估标准单一等。最后研究者们基于决策树模型原理建立了 ICT 模型,除进行了相应报道^[16,24,27-32]外,还实现了 ICT 相应软件“暴力风险分级(classification of violence risk, COVR)”的开发^[33]。

ICT 实际上是数个标准决策树模型(如 CHAID 或

CART)的重复组合。该模型确定了两个临界点,分别用以判断低风险(低于低临界点者)和高风险(高于高临界点者)对象,此后,对介于高和低风险之间的被试(患者)再次采用新的分类树模型作进一步分析^[31]。在以上研究当中,以暴力风险程度(高风险、低风险和低风险不确定)作为应变量,预测将发生暴力的概率高于所研究目标群体暴力基线率 2 倍的个体被归入“高风险”组,低于暴力基线率 0.5 倍的个体则被归入“低风险”组,而被预测概率介于两者之间的个体将被定义为风险不确定或“中等风险”组。基于第一个标准决策树模型的预测结果,第二个标准决策树模型将针对第一个模型结果中的“中等风险”组作进一步分类预测(再寻找“高风险”、“低风险”组),此过程反复进行,直至不能将更多的个体分入“高风险”、“低风险”组(树终节点内的个体数不低于 50 例)^[30]。ICT 模型中被报道最高精度为 AUC 值可达 0.88^[29],但是这些研究普遍存在一个问题,由于现实中缺少对“高风险”、“低风险”在个体水平上的金标准,因此,以上研究缺乏与实际情况相关的金标准对其的验证。因此,对 ICT 模型的验证和推广仍有待进一步研究。此外,有学者质疑这种复杂的分类系统并不比传统的回归方法简单^[26]。

3 神经网络模型

神经网络模型来自对人工智能的研究,其最大特点一是对容错性的模仿,二是通过构造低水平的脑结构以实现生物神经网络的学习能力^[34]。此外,还具有对数据分布特征要求低、可以构建几乎各种非线性结构等优点^[35]。目前,该模型常被用于解决许多计算量庞大的问题,以辨识多重输入变量之间不为人脑所识别的复杂关系和模式^[36]。神经网络的种类较多,如多层感知器网络(multilayer perceptron networks)、概率神经网络(probabilistic neural networks)、广义回归神经网络(general regression neural networks)、径向基函数神经网络(radial basis function networks)、级联算法(cascade correlation)、函数链接型神经网络(functional link networks)、学习矢量量化网络(learning vector quantization networks)、循环神经网络(recurrent networks)、混合网络(hybrid networks)、Kohonen 网络(Kohonen networks)、Gram-Charlier 网络(Gram-Charlier networks)、Hebb 网络(Hebb networks)、自适应神经元网络(Adaline networks)、异联想网络(Heteroassociative networks)等^[37]。

多层感知器(multilayer perception,MLP)是目前神经网络中发展最为成熟和应用最为广泛的一种模型^[38]。如未作特别注明,大多数文献中所指的神经网络

或人工神经网络(artificial neural networks,ANN)实际上都是指 MLP^[37],该模型较为稳定且有较好的容错能力。其前向反馈结构(图 2)包括:输入层(input layer)、隐含层(hidden layer)、输出层(output layer)、偏倚项(bias term)以及权重系数(weight)^[39]。MLP 中的隐含层数目和各隐含层中的隐含单位(hidden units)数目越多,其模拟和运算功能则越强大^[34]。以 MLP 为代表的神经网络的预测功能是通过训练学习(train)和测试(test)来实现。其训练学习过程包括:(1)输入一对“自变量-目标变量”(input-target pair);(2)计算相应预测值;(3)根据误差函数(error function)计算预测值与对应目标值间的差异;(4)采用训练算法(training algorithm)调整网络内部的权重系数;第(1)至(4)步的训练循环(training cycle)称为一次迭代(iteration);(5)不断重复训练迭代,直到模型有效产生较为准确的预测值^[40]。许多不同领域的研究均显示,与传统统计方法相比较,神经网络模型在分类预测上的预测效果与传统方法相当,甚至更优^[41-44]。

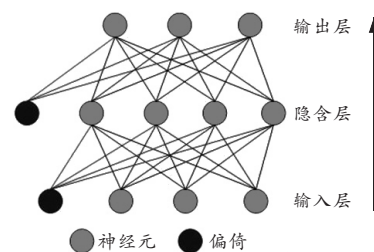


图 2 MLP 结构示意图

影响暴力发生的各种因素之间的关系和模式是复杂且难以识别的。犯罪的影响因素众多,如个体(内在的)因素(如年龄、性别、婚姻状态、精神疾病有无等),环境(外在的)因素(如暴露于危险环境中、是否受外界刺激等)。同时,随时间改变,各种因素可能交互影响,这些都决定了犯罪预测模式的建立将非常复杂和困难。而且,影响因素的数据分布特征往往不符合一些传统统计方法对数据的要求。因此,用神经网络模型来进行风险预测应该具有一定的潜力^[45]。

Brodzinski 等^[46]首次将神经网络模型用于再犯分类预测。他们对美国 778 名青少年缓刑犯的资料进行了分析,将样本划分为训练样本 390 例、测试样本 388 例,结果在测试样本中得到了高达 99%的分类准确度,而与之相比较的判别分析则只达到了 66%的总体准确度。但由于该文献未给出神经网络模型在训练样本中的分类准确度,尚无法判断该研究中是否存在过度拟合(overfitting)的问题。Caulkins 等^[41]对神经网络模型和传统统计方法(包括多重回归、关联分析等)的再犯预测效果进行了比较,采用回顾性方法收

集了1970至1972年从美国监狱释放的3389名犯人的随访资料,采用反映当前犯罪、历史犯罪、社会情况和监狱改造情况等29个变量做预测因子,结果显示神经网络模型的预测效果并没有较传统统计方法更优。由于文中的几种预测方法所得预测准确性无论在训练和测试样本中均较差,研究者们总结认为导致预测失败的原因并非所选统计方法预测效能不足,而有可能是由于预测因子选择不当等因素造成。2000年,Palocsay等^[35]比较了神经网络模型和Logistic回归模型对再犯预测的效果,收集了1978至1980年从美国北卡罗来纳州监狱释放出的超过10000名犯人的随访资料。预测因子共9个,包括性别、是否为非洲裔美国人、是否有严重的酗酒史、是否有烈性毒品使用史等。结果显示两种方法的分类正确率均在60%~69%,而神经网络模型略佳,且差异有统计学意义。研究者们认为神经网络模型的表现虽然较大程度上依赖于其网络拓扑结构(包括隐含层数量、隐含层内的隐含神经元数量、激活函数种类)的选择,但却可能在再犯预测领域中比传统统计方法更具优势,原因是神经网络模型应用起来更灵活、适用范围更广等。之后,Grann等^[47]将神经网络模型、传统的Nuffield法、Logistic回归模型和非加权法(即直接把预测因子得分简单相加进行ROC分析)的暴力再犯预测效果进行了比较,他们使用历史-临床-风险-20项清单(historical-clinical-risk management-20, HCR-20)对瑞典的404名司法精神病患者进行评估,采用HCR-20中的10项历史条目作为预测因子。结果显示,各方法的AUC变化范围在0.51~0.79,训练样本中Logistic回归模型和神经网络模型均未比非加权法或Nuffield法更准确;而在验证样本中,Logistic回归模型和神经网络模型,尤其是神经网络模型的预测效果却明显降低。研究者们认为要想提高预测准确度,预测方法并非是最重要的,首当其冲应该是预测因子的选择和组合。同时,他们还指出神经网络模型的最大问题在于其“黑箱”特征,这种可解释性差的模型对临床和实践是没有太大帮助的。

针对以上3类方法在风险评估中孰优孰劣所表现出的不同效果和报道,Liu等^[48]将Logistic回归、CART和MLP3种模型同时应用于暴力风险评估并进行了比较。该研究对英国1225名男性监狱犯人释放后是否发生暴力再犯进行了为期平均3.31年的随访,以HCR-20的20项条目得分作为预测因子对暴力再犯这一结局变量进行了预测评估。在经过多重验证和比较后,该研究结果显示3种模型的正确率范围为0.59~0.67,AUC范围为0.65~0.72,神经网络模型

的预测效果总体比另外两种模型略优,但差异没有统计学意义。

4 结 语

综上,不同的分类预测方法各有优劣。Logistic回归模型在样本足够时表现稳定,且相对最不容易发生过度拟合,但对极端值和特殊观测对象亚组不太敏感;CART等决策树模型能直观生动地反映临床决策过程,能敏感地发现具有不同特征的群组亚类,有利于风险管理和决策,但其稳定性问题尚待进一步研究;MLP作为人工智能技术神经网络模型的代表,对数据和样本含量的要求均较低,在经过充分训练之后,对越是复杂的数据,其学习和数据挖掘能力表现越为突出,并且,在采取适当控制措施并制定“最优”模型判断标准后,MLP也能较好地克服过度拟合问题。但由于“黑箱”原理特质,目前仍因为模型的可解释性和泛化性差而尚未被临床工作者所广泛接受^[49]。恰当的统计预测方法只是暴力风险有效评估的重要环节之一。正如Yang等^[50]在经过系列实证研究之后所提出的决定风险评估效果最为重要且相互制约的四项因素:良好的预测因子、特定的同质目标人群、足够区分度的因变量亚类水平和恰当的统计预测方法。未来风险评估可考虑这样一种发展模式,多步骤、分阶段地完成。

参考文献:

- [1] Hanson RK. Twenty years of progress in violence risk assessment[J]. *J Interpers Violence*, 2005, 20(2): 212-217.
- [2] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve[J]. *Radiology*, 1982, 143(1): 29-36.
- [3] Nuffield J. Parole decision-making in Canada: Research towards decision guidelines[M]. Ottawa: Communication Division, Solicitor General of Canada, 1982.
- [4] Kirby BC. Parole prediction using multiple correlation[J]. *American Journal of Sociology*, 1954, 59(6): 539-550.
- [5] van Alstyne DJ, Gottfredson MR. A multidimensional contingency table analysis of parole outcome: New methods and old problems in criminological prediction[J]. *Journal of Research in Crime and Delinquency*, 1978, 15(2): 172-193.
- [6] Borum R. Improving the clinical practice of violence risk assessment. Technology, guidelines, and training[J]. *Am Psychol*, 1996, 51(9): 945-956.
- [7] Cooper BS, Griesel D, Yuille JC. Clinical-forensic risk assessment: The past and current state of affairs[J].

- J Forensic Psychol Pract, 2008, 7(4): 1-63.
- [8] Gottfredson SD, Moriarty LJ. Statistical risk assessment: Old problems and new applications[J]. *Crime & Delinquency*, 2006, 52(1): 178-200.
- [9] Hosmer DW, Lemeshow S. Applied logistic regression[M]. 2nd ed. New York: John Wiley & Sons, 2000:1.
- [10] Harper PR. A review and comparison of classification algorithms for medical decision making[J]. *Health Policy*, 2005, 71(3): 315-331.
- [11] Albert A, Anderson JA. On the existence of maximum likelihood estimates in logistic regression models[J]. *Biometrika*, 1984, 71(1): 1-10.
- [12] Gardner W, Lidz CW, Mulvey EP, *et al.* A comparison of actuarial methods for identifying repetitively violent patients with mental illnesses[J]. *Law Hum Behav*, 1996, 20(1): 35-48.
- [13] Thomas S, Leese M, Walsh E, *et al.* A comparison of statistical models in predicting violence in psychotic illness[J]. *Compr Psychiatry*, 2005, 46(4): 296-303.
- [14] Hartvig P, Alfarnes S, Ostberg B, *et al.* Brief checklists for assessing violence risk among patients discharged from acute psychiatric facilities: A preliminary study[J]. *Nord J Psychiatry*, 2006, 60(3): 243-248.
- [15] 王小平, 杨德森, 李凌江, 等. 精神分裂症患者攻击行为的预测[J]. *中华精神科杂志*, 1997, 30(1): 20-24.
- [16] Steadman HJ. From dangerousness to risk assessment of community violence: Taking stock at the turn of the century[J]. *J Am Acad Psychiatry Law*, 2000, 28(3): 265-271.
- [17] Morgan JN, Sonquist JA. Problems in the analysis of survey data and a proposal[J]. *J Am Stat Assoc*, 1963, 58(302): 415-434.
- [18] Breiman L, Friedman JH, Olshen RA, *et al.* Classification and regression trees[M]. Boca Raton: Chapman & Hall/CRC, 1998.
- [19] Kass GV. An exploratory technique for investigating large quantities of categorical data[J]. *Appl Statist*, 1980, 29(2): 119-127.
- [20] Quinlan JR. C4.5: Programs for machine learning[M]. San Mateo: Morgan Kaufmann Publishers, 1993.
- [21] Quinlan JR. Improved use of continuous attributes in C4.5[J]. *J Artif Intell Res*, 1996, 4: 77-90.
- [22] Loh WY, Shih YS. Split selection methods for classification trees[J]. *Statistica Sinica*, 1997, 7: 815-840.
- [23] Friedman JH. Greedy function approximation: A gradient boosting machine[J]. *The Annals of Statistics*, 2001, 29(5): 1189-1232.
- [24] Steadman HJ, Mulvey EP, Monahan J, *et al.* Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods[J]. *Arch Gen Psychiatry*, 1998, 55(5): 393-401.
- [25] Stalans LJ, Yarnold PR, Seng M, *et al.* Identifying three types of violent offenders and predicting violent recidivism while on probation: A classification tree analysis[J]. *Law Hum Behav*, 2004, 28(3): 253-271.
- [26] Rosenfeld B, Lewis C. Assessing violence risk in stalking cases: A regression tree approach[J]. *Law Hum Behav*, 2005, 29(3): 343-357.
- [27] Silver E, Smith WR, Banks S. Constructing actuarial devices for predicting recidivism: A comparison of methods[J]. *Crim Justice Behav*, 2000, 27(6): 733-764.
- [28] Silver E, Chow-Martin L. A multiple models approach to assessing recidivism risk: Implications for judicial decision making[J]. *Crim Justice Behav*, 2002, 29(5): 538-568.
- [29] Banks S, Robbins PC, Silver E, *et al.* A multiple-models approach to violence risk assessment among people with mental disorder[J]. *Crim Justice Behav*, 2004, 31(3): 324-340.
- [30] Monahan J, Steadman HJ, Appelbaum PS, *et al.* Developing a clinically useful actuarial tool for assessing violence risk[J]. *Br J Psychiatry*, 2000, 176: 312-319.
- [31] Monahan J, Steadman HJ, Silver E, *et al.* Rethinking risk assessment: The MacArthur study of mental disorder and violence[M]. New York: Oxford University Press, 2001.
- [32] Monahan J, Steadman HJ, Robbins PC, *et al.* An actuarial model of violence risk assessment for persons with mental disorders[J]. *Psychiatr Serv*, 2005, 56(7): 810-815.
- [33] Monahan J, Steadman HJ, Appelbaum PS, *et al.* The classification of violence risk[J]. *Behavioral Sciences & the Law*, 2006, 24(6): 721-730.
- [34] Cheng B, Titterton DM. Neural networks: A review from a statistical perspective[J]. *Statistical Science*, 1994, 9(1): 2-54.
- [35] Palocsay SW, Wang P, Brookshire RG. Predicting criminal recidivism using neural networks[J]. *Socio-Economic Planning Sciences*, 2000, 34(4): 271-284.
- [36] Bigi R, Gregori D, Cortigiani L, *et al.* Artificial neural networks and robust Bayesian classifiers for risk stratification following uncomplicated myocardial infarction[J]. *Int J Cardiol*, 2005, 101(3): 481-487.
- [37] Software for predictive modeling and forecasting: Multilayer perceptron neural networks[Z/OL]. [2012-08-20]. <http://www.dtreg.com/mlfn.htm>.
- [38] Price RK, Spitznagel EL, Downey TJ, *et al.* Applying artificial neural network models to clinical decision

- making[J]. *Psychol Assess*,2000,12(1):40-51.
- [39] Bishop CM. *Neural networks for pattern recognition*[M]. New York: Oxford University Press,1995.
- [40] Model extremely complex functions, neural networks[Z/OL]. [2012-08-20]. <http://www.statsoft.com/textbook/neural-networks/#multilayera>.
- [41] Caulkins J, Cohen J, Gorr W, *et al*. Predicting criminal recidivism: A comparison of neural network models with statistical methods[J]. *J Crim Justice*, 1996,24(3):227-240.
- [42] Liang T, Chandler JS, Han I, *et al*. An empirical investigation of some data effects on the classification accuracy of probit, ID3, and neural networks[J]. *Contemporary Accounting Research*,1992,9(1):306-328.
- [43] Sharda R. Neural networks for the MS/OR analyst: An application bibliography[J]. *Interfaces*,1994,24(2):116-130.
- [44] Subramanian V, Hung MS, Hu MY. An experimental evaluation of neural networks for classification[J]. *Computers & Operations Research*,1993,20(7):769-782.
- [45] Florio T, Einfeld S, Levy F. Neural networks and psychiatry: candidate applications in clinical decision making[J]. *Aust N Z J Psychiatry*,1994,28(4):651-666.
- [46] Brodzinski JD, Crable EA, Scherer RF. Using artificial intelligence to model juvenile recidivism patterns[J]. *Computers in Human Services*,1994,10(4):1-18.
- [47] Grann M, Långström N. Actuarial assessment of violence risk: To weigh or not to weigh?[J]. *Crim Justice Behav*,2007,34(1):22-36.
- [48] Liu YY, Yang M, Ramsay M, *et al*. A comparison of logistic regression, classification and regression tree, and neural networks models in predicting violent re-offending[J]. *J Quant Criminol*,2011,27(4):547-573.
- [49] 刘元元. 暴力再犯风险分类预测方法比较研究及其非平衡数据问题探讨[D].成都:四川大学华西公共卫生学院,2010:99.
- [50] Yang M, Liu Y, Coid J. Applying Neural Networks and other statistical models to the classification of serious offenders and the prediction of recidivism[R]. Ministry of Justice Research Series 6/10,2010:1-30.

(收稿日期:2012-08-21)

(本文编辑:张钦廷)

(上接第215页)

- [11] 朱传红,史绍杏,王海生,等. Y-STR家系排查法的应用原则及注意事项[J]. *中国法医学杂志*,2007,22(6):431-432.
- [12] 史绍杏,马伟,朱传红. Y-STR家系分型及 ITO 分析法侦破强奸杀人案 1 例[J]. *刑事技术*,2005,(4):58-59.
- [13] STR-DNA Y-chromosome[DB/OL]. [2012-10-18]. http://www.denverda.org/DNA/Y-chromosome_DNA_Lega_%20Decisions.htm.
- [14] Ge J, Eisenberg A, Budowle B. Developing criteria and data to determine best options for expanding the core CODIS loci[J]. *Investig Genet*,2012,3:1.
- [15] Davis C, Ge J, Sprecher C, *et al*. Prototype PowerPlex® Y23 System: A concordance study[J]. *Forensic Sci Int Genet*,2013,7(1):204-208.
- [16] Ge J, Budowle B, Aranda XG, *et al*. Mutation rates at Y chromosome short tandem repeats in Texas populations[J]. *Forensic Sci Int Genet*,2009,3(3):179-184.
- [17] 史绍杏,朱传红,王海生,等. Y-STR家系排查中出现 Y-STR 突变 1 例[J]. *刑事技术*,2008,(1):68-69.
- [18] Ge J, Budowle B, Chakraborty R. DNA identification by pedigree likelihood ratio accommodating population substructure and mutations[J]. *Investig Genet*, 2010,1(1):8.

(收稿日期:2012-10-19)

(本文编辑:李莉)