机器学习算法在预测男男性行为人群中 HIV 感染的应用*

天津医科大学公共卫生学院流行病与卫生统计学系(300070)
郭长满　郭　敏　刘媛媛　李长平　崔　壮△　马　骏

【提　要】 目的　应用不同机器学习算法预测男男性行为(MSM)人群 HIV 感染状况的比较。方法　将四种机器学习算法(logistic 回归、神经网络、随机森林和支持向量机)的预测结果和实验室的检测结果相比较,分类性能的评价指标采用 ROC 曲线下面积(AUC)、灵敏度、特异度和准确度(PRE)。结果　四种算法在训练集和测试集上均具有较理想的分类效能,训练集的结果略好于测试集。和 logistic 回归分析相比较,其他几种算法均提高了分类预测效能:神经网络提高 18.4%(*AUC*:0.909,95% *CI*:0.903~0.915),随机森林提高 19.7%(*AUC*:0.922,95% *CI*:0.920~0.924),支持向量机提高 22.3%(*AUC*:0.948,95% *CI*:0.947~0.949)。其中支持向量机的分类性能最好,分类的灵敏度为 97.5%,特异度为 99.1%,准确度(PRE)为 98.9%。结论　机器学习算法显著地提高了 MSM 人群中 HIV 感染的预测效能,可以较准确地识别 MSM 人群中 HIV 感染者与未感染者,为及时地提供预防与治疗服务提供了依据,同时避免了医疗资源的浪费。

【关键词】 机器学习算法　男男性行为人群　HIV

Application of Machine-learning Algorithms in Predicting HIV Infection among Men Who Have Sex with Men

Guo Changman, Guo Min, Liu Yuanyuan, et al(*Department of Health Statistics, Public Health College, Tianjin Medical University*(300070), *Tianjin*)

【Abstract】 Objective　A comparative study of machine-learning algorithms on predicting HIV infection among men who have sex with men (MSM). **Methods**　Comparing the predicted results of four machine-learning algorithms (logistic Regression, Neural Network, Random Forests and Support Vector Machines) with the laboratory test results, the evaluation of the classified prediction results is evaluated by Area under the Curve (AUC), sensitivity, specificity and accuracy (PRE). **Results**　All the four algorithms have ideal classification efficiency, and the result of the training set is slightly better than that of the test set. Compared to the logistic regression, other algorithms have improved prediction: Neural Network rose 18.4% (AUC: 0.909, 95% CI: 0.903-0.915), Random Forest rose 19.7% (AUC: 0.922, 95% CI: 0.920-0.924), Support Vector Machines rose 22.3% (AUC: 0.948, 95% CI: 0.947-0.949), the best classification algorithm is Support Vector Machines (sensitivity 97.5%, specificity 99.1%, accuracy 98.9%). **Conclusion**　The machine-learning algorithms significantly improved the predictive effectiveness of HIV infection among men who have sex with men, which successfully identified HIV infected MSM and provided the basis for the timely provision of prevention and treatment services.

【Key words】 Machine-learning algorithms; MSM; HIV

男男性行为人群(men who have sex with men, MSM)是感染 HIV、性病风险最高的人群之一,也是感染人数增长较快的人群[1],在最新确认的 HIV 感染者中,MSM 所占比例稳步增长,甚至一度达到了新确诊的 22.8%[2]。当前的研究已经表明高危性行为,如多性伴、群交、使用物质(助性剂)以及无保护性交均为 HIV 感染的高危因素[3]。针对具有这些高危因素的人群采取必要的预防保护措施可以有效地减少 HIV 在该人群中的传播,提高该人群的健康水平,尽管当前已经有大量的检测措施可以早期发现和治疗 HIV 感染者,但是每年仍然有大量未被发现的新增 HIV 感染者,并且有一部分人群仍在接受不必要的预防服务,从而造成医疗资源的浪费,因此,开发一种准确而有效的识别早期 HIV 感染者的方法,具有重要的现实意义。

已有的模型如 logistic 回归分析和 Poisson 回归分析已经在男男性行为人群中的 HIV 感染广泛应用,然而这些模型在男男性行为人群中的分类和预测性能却少有研究,机器学习算法的发展为评估该高危人群的特征提供了一种新的思路。

机器学习又称为人工智能,即通过计算机网络处理各个变量间的复杂和非线性关系并使误差最小化的方法[4]。目前广泛应用的机器学习算法包括神经网络、随机森林和支持向量机,这些算法已经广泛应用于工程学、建筑学等领域,却很少有研究将这些算法应用于男男性行为人群,为了更好地评估这些算法是否能提高预测 HIV 感染的精确度,以及寻找具有最好分类效能的分类算法,本研究比较了四种算法的分类效能。

原理与方法

1. logistic 回归的原理

logistic 回归分析在医学研究中应用广泛。目前主要是用于流行病学研究中危险因素的筛选,但它同

* 基金项目:教育部人文社会科学研究项目(11YJCZH022);中国性病艾滋病防治协会高校防艾基金(2017120101B000359)
△通信作者:崔壮,E-mail:cuizhuang@tmu.edu.cn

时具有良好的判别和预测功能,尤其是在资料类型不能满足 Fisher 判别和 Bayes 判别的条件时,更显示出 logistic 回归判别的优势和效能[5]。

2. BP 神经网络的原理

BP 神经网络是一种有监督的前馈运行的神经网络,它由输入层、隐含层、输出层以及各层之间的节点的连接权所组成,这个学习过程的算法由信息的正向传播和误差的反向传播构成,在正向传播过程中,输入信息从输入层经隐含层逐层处理,并传向输出层,每一层神经元只影响下一层神经元的输出,信息完成正向的传播后,如果在输出层不能得到期望的输出,那么误差将进入反向传播,运用链导数法则将连接权关于误差函数的导数沿原来的连接通路返回,通过修改各层的权值使得误差函数减小[6]。

3. 随机森林的原理

随机森林由 Leo Breiman(2001)提出,它通过自助法(bootstrap)重采样技术,从原始训练样本集 N 中有放回地重复随机抽取 n 个样本生成新的训练自助样本集合,然后根据自助样本集生成 n 个分类树组成随机森林,新数据的分类结果按分类树投票多少形成的分数而定[7]。

4. 支持向量机的原理

支持向量机通过结构风险最小化原理来提高泛化能力,它较好地解决了小样本、非线性、高维数、局部极小点等实际问题。其主要思想:首先选择一非线性映射把 n 维样本从原空间映射到特征空间,在此高维特征空间中构造最优线性决策函数。在构造最优决策函数时,利用了结构风险最小化原则,同时引入了间隔的概念。并巧妙地利用原空间的核函数取代了高维特征空间的点积运算,避免了复杂计算[8]。

5. 算法的比较

本研究纳入了四种常用的数据分类算法,即 logistic 回归、神经网络、随机森林和支持向量机,比较这四种分类算法基于已有的变量信息对目标人群是否感染 HIV 进行分类。为了比较四种分类算法的分类效果,将数据集分为训练集和测试集,训练集用于对分类算法进行训练,测试集用于对训练的结果进行比较和总结。原数据集分别经过 10 次、50 次和 100 次有放回 bootstrap 重抽样[9],从而产生 10 个、50 个和 100 个与原数据集大小相同的子样本集,基于 bootstrap 重抽样的特性,每次抽样时原数据集中总会有约 37% 的样本不被抽到,用这部分不被抽到的样本集来分别作为测试集,新产生的子样本集来分别作为训练集,基于每种分类算法的分类结果进行综合评价。

6. 统计学方法

分类器的分类性能采用测试集的分类结果来进行评价,分类效果的评价采用 C 统计量来进行[10],即曲线下面积(AUC),及其 95% 置信区间,用实验室检测得到的样本人群 HIV 感染情况作为金标准,而每个分类器每次采用验证集分类的结果和金标准进行比较从而可以得到灵敏度、特异度、精确度和相应的曲线下面积。关于神经网络、支持向量机和随机森林最优参数的选取基于 3 折交叉验证的方法,最优模型的选取依据分类模型的曲线下面积,选择具有最大曲线下面积时所对应的参数。其中,神经网络的隐藏层神经元个数范围为(0,10),支持向量机选择的核函数为径向基核函数,对于 cost 设置参数选择范围为(2^{-5},2^{0},2^{15}),gamma 的范围为(2^{-15},2^{0},2^{3}),随机森林中节点数范围为(3,4,5),决策树的个数为范围为(100,200,500),从中选择最佳的参数来进行建模和预测。Nnet 包被用来实现神经网络算法,randomForest 包用来实现随机森林算法,e1071 包用来实现支持向量机算法,rminer 包用于模型调参。所有的统计分析均运用 R 语言实现的。

结　果

1. 研究人群和研究变量

本次研究的资料来源于天津市某男性同性恋志愿组织调查收集的关于男男性行为人群的资料和体检信息,入选标准:①年龄≥18 周岁;②在天津市居住≥6 个月;③在过去六个月曾发生过至少一次商业男男性行为。对数据进行核查、清洗,排除不符合入选标准,数据大量缺失以及有逻辑错误的样本。最终纳入研究的目标人群有 3086 人。对研究变量与 HIV 的关系进行单因素分析,筛选出结果有意义的,以及文献研究显示可能有影响的变量。该目标人群 HIV 感染率为 8.39%。最终研究中用到的变量如表 1 所示。

表 1　研究中纳入的变量

变量	描述
工作地点	个体/单位
工作方式	兼职/全职
性角色	攻为主/受为主
是否参加过群交	是/否
最近一次性交时是否使用安全套	是/否
是否患过性病	是/否
是否做过艾滋病检测	是/否
婚姻状况	未婚/已婚
户籍	本地/外地
性取向	同性恋;异性恋;双性恋
初次性行为年龄	<18 岁;≥18 岁
最近一个月性交人数	≤9 个;>9 个
最近一个月肛交次数	≤10 个;>10 个
月收入	<3000 元;3000~5000 元;>5000 元
联系性伴方式	网络途径;非网络途径
文化程度	初中及以下;高中或中专;大专及以上
持续使用避孕套	是/否
是否使用物质	是/否
从业时长	≤13 个月;>13 个月
是否患 HIV	是/否

2. 分类算法在训练集上的表现

表 2 显示了经过 10 次、50 次和 100 次重抽样后，计算四种分类算法在训练集里的指标及其 95% CI，结果支持向量机在灵敏度、特异度、准确度(PRE) 以及曲线下面积(AUC) 上表现最好。

表 2　四种分类算法在训练集上的分类效能

抽样次数	评价指标	分类算法			
		logistic 回归	神经网络	随机森林	支持向量机
10 次	灵敏度	0. 350(0. 332，0. 368)	0. 958(0. 912，1. 000)	0. 987(0. 979，0. 996)	0. 978(0. 954，1. 000)
	特异度	0. 936(0. 932，0. 940)	0. 977(0. 969，0. 985)	0. 982(0. 980，0. 984)	0. 991(0. 990，0. 991)
	PRE	0. 887(0. 879，0. 895)	0. 975(0. 965，0. 985)	0. 983(0. 982，0. 984)	0. 989(0. 989，0. 989)
	AUC	0. 733(0. 721，0. 745)	0. 907(0. 877，0. 937)	0. 928(0. 920，0. 936)	0. 949(0. 941，0. 957)
50 次	灵敏度	0. 331(0. 319，0. 343)	0. 952(0. 940，0. 964)	0. 984(0. 980，0. 988)	0. 971(0. 953，0. 989)
	特异度	0. 934(0. 933，0. 935)	0. 982(0. 980，0. 984)	0. 982(0. 981，0. 983)	0. 991(0. 990，0. 992)
	PRE	0. 879(0. 875，0. 884)	0. 979(0. 977，0. 981)	0. 982(0. 982，0. 982)	0. 990(0. 989，0. 991)
	AUC	0. 738(0. 734，0. 742)	0. 926(0. 916，0. 936)	0. 929(0. 925，0. 933)	0. 955(0. 949，0. 961)
100 次	灵敏度	0. 331(0. 323，0. 339)	0. 953(0. 947，0. 959)	0. 980(0. 977，0. 983)	0. 985(0. 979，0. 991)
	特异度	0. 936(0. 935，0. 937)	0. 982(0. 981，0. 983)	0. 982(0. 982，0. 982)	0. 990(0. 989，0. 991)
	PRE	0. 881(0. 879，0. 883)	0. 978(0. 977，0. 979)	0. 982(0. 982，0. 982)	0. 990(0. 990，0. 990)
	AUC	0. 740(0. 736，0. 744)	0. 925(0. 919，0. 931)	0. 928(0. 926，0. 930)	0. 948(0. 948，0. 948)

3. 分类算法在测试集上的表现

表 3 显示了经过 10 次、50 次和 100 次重抽样后四种分类算法在测试集上的效能指标及其 95% *CI*，结果显示随机森林的灵敏度最高(97. 6%) ，支持向量机在特异度，准确度(PRE) 以及曲线下面积(AUC) 上表现最好。

表 3　四种分类算法在测试集上的分类效能

抽样次数	评价指标	分类算法			
		logistic 回归	神经网络	随机森林	支持向量机
10 次	灵敏度	0. 327(0. 311，0. 343)	0. 949(0. 929，0. 969)	0. 976(0. 960，0. 992)	0. 973(0. 953，0. 993)
	特异度	0. 937(0. 931，0. 943)	0. 979(0. 958，1. 000)	0. 979(0. 977，0. 981)	0. 990(0. 989，0. 991)
	PRE	0. 886(0. 880，0. 892)	0. 976(0. 971，0. 981)	0. 980(0. 979，0. 981)	0. 989(0. 989，0. 989)
	AUC	0. 724(0. 716，0. 732)	0. 916(0. 876，0. 956)	0. 924(0. 918，0. 930)	0. 949(0. 943，0. 955)
50 次	灵敏度	0. 304(0. 294，0. 314)	0. 931(0. 919，0. 943)	0. 977(0. 971，0. 983)	0. 973(0. 965，0. 981)
	特异度	0. 937(0. 936，0. 938)	0. 978(0. 976，0. 980)	0. 981(0. 980，0. 982)	0. 991(0. 991，0. 991)
	PRE	0. 879(0. 875，0. 883)	0. 973(0. 971，0. 975)	0. 980(0. 979，0. 981)	0. 990(0. 990，0. 990)
	AUC	0. 723(0. 721，0. 725)	0. 908(0. 898，0. 918)	0. 921(0. 919，0. 923)	0. 949(0. 947，0. 951)
100 次	灵敏度	0. 306(0. 300，0. 312)	0. 931(0. 923，0. 939)	0. 976(0. 972，0. 980)	0. 975(0. 969，0. 981)
	特异度	0. 937(0. 936，0. 938)	0. 978(0. 977，0. 979)	0. 981(0. 981，0. 981)	0. 991(0. 991，0. 991)
	PRE	0. 879(0. 837，0. 881)	0. 973(0. 972，0. 974)	0. 979(0. 979，0. 979)	0. 989(0. 989，0. 989)
	AUC	0. 725(0. 724，0. 726)	0. 909(0. 903，0. 915)	0. 922(0. 920，0. 924)	0. 948(0. 947，0. 949)

4. 四种不同分类算法预测性能比较

预测性能用曲线下面积(AUC) 来表示，分别经过 10 次、50 次和 100 次 bootstrap 重抽样后: logistic 回归分类结果对应的 *AUC* 分别是为 0. 724、0. 723 和 0. 725; 神经网络为 0. 916、0. 908 和 0. 909; 随机森林为 0. 924、0. 921 和 0. 922; 支持向量机为 0. 949、0. 949 和 0. 948; 经过 100 次重抽样后，相比于 logistic 回归，神经网络、随机森林和支持向量机的预测性能分别提升了 18. 4%、19. 7% 和 22. 3%，具体可参见表 4。

表 4　不同分类算法预测男性同性恋人群 HIV 的比较

抽样次数	分类算法	*AUC*	标准误	95% *CI*		提升度
				下限	上限	
10 次	logistic 回归	0. 724	0. 004	0. 716	0. 732	–
	神经网络	0. 916	0. 020	0. 876	0. 956	0. 192
	随机森林	0. 924	0. 003	0. 918	0. 930	0. 200
	支持向量机	0. 949	0. 004	0. 943	0. 955	0. 223
50 次	logistic 回归	0. 723	0. 002	0. 721	0. 725	–
	神经网络	0. 908	0. 050	0. 898	0. 918	0. 185
	随机森林	0. 921	0. 002	0. 919	0. 923	0. 198
	支持向量机	0. 949	0. 001	0. 947	0. 951	0. 226
100 次	logistic 回归	0. 725	0. 001	0. 724	0. 726	–
	神经网络	0. 909	0. 003	0. 903	0. 915	0. 184
	随机森林	0. 922	0. 001	0. 920	0. 924	0. 197
	支持向量机	0. 948	0. 0005	0. 947	0. 949	0. 223

5. 变量重要性

图 1 列出了所有变量的重要性，并使用训练集进行计算，通过设置各种算法的最优参数得到每种算法训练 100 次后变量的平均重要性。HIV 感染的预测算法的变量重要性列于图 1。

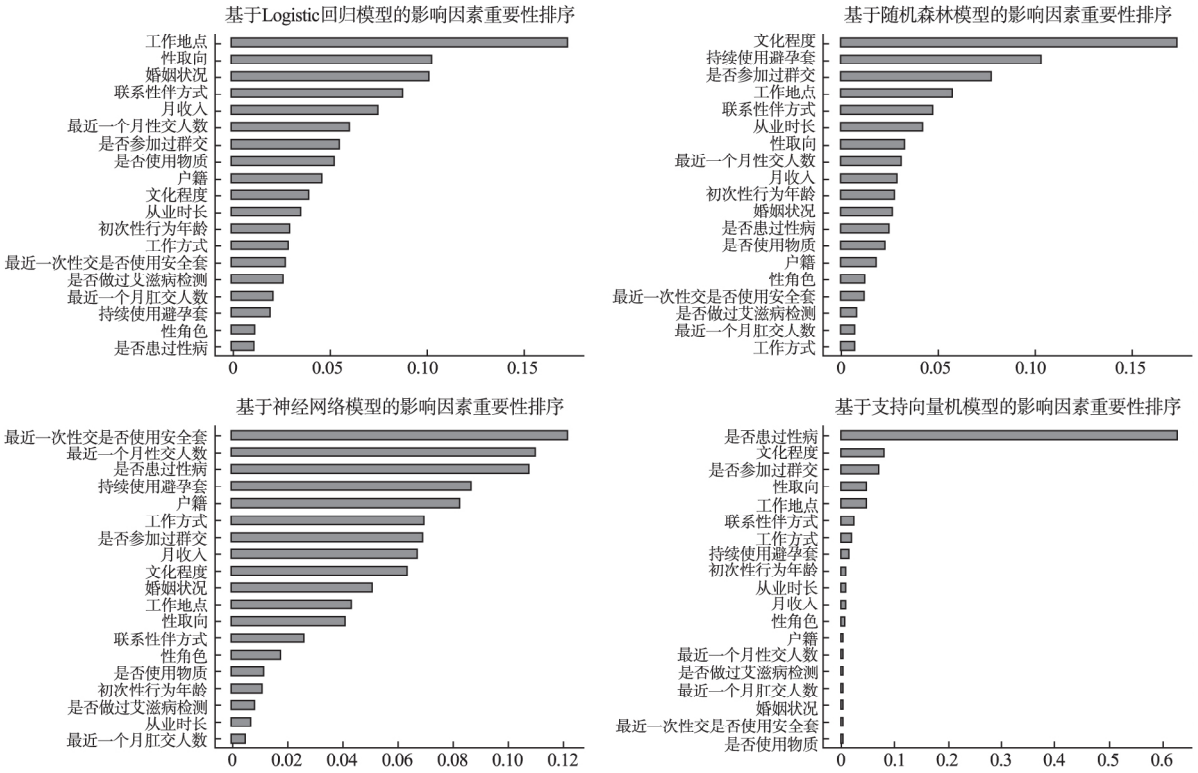

图 1　基于机器学习算法的变量重要性

讨　论

男男性行为人群是 HIV 感染的高危人群[3]，近年来的研究发现该人群 HIV 感染率在 10% 左右[11]。一方面由于其隐蔽性和不可及性，该人群的健康状况资料较难获得，因此研究该目标人群时常常受到样本量的限制。另一方面，由于 HIV 的高危性，一旦感染 HIV 却未得到及时的抗病毒治疗将会导致病情的发展并最终导致艾滋病的发生。因此为了实现基于有限的样本数据，提高分类模型（或分类器）的分类能力的目标，我们尝试采用传统模型和机器学习算法结合，探索适用于男男性行为人群研究的最佳分类器。

本研究是第一次将机器学习算法应用到男男性行为人群中进行分类的研究，采用 bootstrap 抽样方法用来对数据进行抽样，结果具有较好的稳定性[9]，经过 100 次 bootstrap 重抽样后，相比于 logistic 回归，神经网络、随机森林和支持向量机表现出较好的分类效能，所对应的曲线下面积（AUC）分别提高了 18.4%、19.7% 和 22.3%，且支持向量机为最优分类算法，有最高的分类准确度（98.9%）和曲线下面积（94.8%）。

在变量的重要性的计算中，不同的算法具有不同的理论基础[12]。其中 logistic 回归的变量重要性用回归系数和标准差的乘积来衡量。随机森林是基于平均基尼系数或平均精确度减少量；神经网络使用模型内变量的总体加权；支持向量机则是基于信息值的变化来衡量。结果显示各个变量在不同算法中的重要性大小不一，但是综合上述算法最终结果显示，高危性行为及性病史仍然是影响 HIV 感染的主要因素，这与之前的研究结果相一致[3]，因此洁身自好，养成良好的生活方式仍然是预防 HIV 感染的关键措施。

随着潜在风险因素的数量增加，模型的复杂性可能导致过度拟合，产生不可信的结果。为了避免该问题，常用的方法包括适当选择预训练、调整超参数、交叉验证、bootstrap 和正则化等[13]。本研究中，我们通过对原数据集进行 bootstrap 重抽样并对结果进行 10 次、50 次和 100 次的循环来验证用训练样本训练的模型稳定性，对比训练集和测试集的结果显示，两者差异不大，分类效能均比较理想，表明模型的泛化能力比较好；其次通过对机器学习算法中超参数的调整，选择最优的超参数使模型达到最优的分类效能。

作为经典的统计学方法，logistic 回归仍然是一个可靠的分类方法，其可以计算出各个变量在模型中的系数以及优势比，各个变量在模型中的作用是清晰、明确的。但是对于非线性可分问题，或处理分类能力有限的变量时表现往往不佳。机器学习算法如神经网络、支持向量机和随机森林已成为统计学研究的热点，因其具有较强自适应、自学习、非线性映射、容错和泛化能力，正在越来越多地被应用到实际问题中。应用神经网络时如何选取合适的隐藏层是其中的关键[6]，本研究中选取每次训练结果（AUC）最好时的参数作为每次测试集的最优参数。支持向量机算法在处理高维小样本数据时具有比较好的分类效能。其最优模型参数的选取是基于每次训练过程中模型最优性能时所

（下转第 35 页）

合标准。条目 COPD1(您咳嗽吗)在难度等级为 B4 时对应的难度系数为(3.29 > 2.95),条目 COPD2(您早晨起床时咳嗽较白天多吗)在难度等级为 B1 时对应的难度系数为(-3.09 < -2.95),均接近标准值,提示可增加样本量来减小误差,若是仍不符合标准,则需要进一步分析原因或者删除条目。

综上所述,MHIEC-COPD 量表条目针对老年慢阻肺患者而制订,共性模块用于老年慢性病患者共性特征体现,特异模块用于疾病状态的测量;共性模块结合特异模块一起使用,因此特异模块条目不宜过多。本研究采用 CTT 与 IRT,从宏观与微观角度对条目进行分析;课题组慢阻肺专家对统计结果进行讨论并提出修改意见,这样就避免了单纯依靠统计分析造成的失误,提高了条目的代表性和可靠性。

参 考 文 献

[1] 李敏捷,吕宏梅,罗艳虹,等. 慢性阻塞性肺疾病患者报告临床结局量表的条目筛选. 中国呼吸与危重监护杂志,2016,15(2):105-108.

[2] 高媛,秦军. 生物燃料烟雾与慢性阻塞性肺疾病研究进展. 临床肺科杂志,2011,16(5):746-747.

[3] De Rossi Figueiredo D,Paes LG,Warmling AM,et al. Multidimensional measures validated for home health needs of older persons: A systematic review. Int J Nurs Stud,2017,25(77):130-137.

[4] 杨铮,李晓梅,万崇华,等. 慢性阻塞性肺疾病患者生命质量测定量表的研制与考评. 中国全科医学,2007,10(13):1080-1083.

[5] 万崇华. 慢性病患者生命质量测评与应用. 北京:科学出版社,2015:81-93.

[6] Embretson SE,Reise SP. Item response theory for psychol-ogists. Mahwah:Lawrence Erlbaum,2000:13-125.

[7] 石志红,曾宪华,罗艳虹,等. IRT 等级反应模型在慢性呼吸衰竭 PRO 量表编制中的应用. 数理医药学杂志,2014,7(4):453-455.

[8] 臧运洪,赵守盈,陈维,等. 用项目反应理论修订父母同伴依恋量表. 贵州师范大学学报:自然科学版,2012,30(2):22-27.

[9] 涂冬波,蔡艳. 信息函数在标准参照测验中的应用研究. 江西师范大学学报:自然科学版,2005,29(2):167-172.

[10] 刘炳伦,郝伟,杨德森,等. 网络依赖诊断量表初步编制. 中国临床心理学杂志,2006,14(3):227-232.

[11] 孙晓敏,关丹丹. 经典测量理论与项目反应理论的比较研究. 中国考试(研究版),2009,(9):10-17.

[12] 谢洋,王佳佳. 项目反应理论在呼吸疾病生存质量研究中的应用. 中国老年学杂志,2017,37(4):1038-1039.

[13] 杨铮. 癌症患者生命质量测定量表共性模块(V2.0)研制及最小临床有意义差异制定. 广州:南方医科大学,2015.

[14] 林岳卿,张伟涛,方积乾. 项目反应理论在医学量表条目筛选中的应用. 中国医药导报,2014,11(5):155-158.

(责任编辑:张 悦)

(上接第 31 页)

对应的参数,参数的选取采用 3 折交叉验证法。随机森林比较适合处理海量数据、高维问题、连续性变量,分类变量等。随机森林在生成过程中采用了 bootstrap 方法进行重抽样,生成其内部的训练集和袋外数据,通过袋外数据来测试模型的分类性能,这种基于 Bagging 的思想提升了模型的性能和稳定性[7],但也存在运算量大的局限性。时至今日,机器学习算法的“黑箱”特性仍被诟病,它们不能像 logistic 回归模型那样描述风险因素变量如何相互作用的复杂性以及它们对结果的独立影响,但数据可视化方法有助于对这些模型的理解[14]。

本研究发现机器学习算法有助于识别未被发现的感染 HIV 的男男性行为人群,从而做到早发现、早诊断、早治疗的目的,同时也为机器学习算法应用于医学数据开辟了思路。

参 考 文 献

[1] 季顺锋. 苏州市 2013-2014 年度男性同性恋性传播性疾病检测结果分析. 中国卫生产业,2016,13(20):120-122.

[2] Wu J. HIV/STIs related risk among middle aged and old MSM in Shenzhen,China. Dissertations & Theses - Gradworks,2016,

[3] Das A,Li J,Zhong F,et al. Factors associated with HIV and syphilis co-infection among men who have sex with men in seven Chinese cities. International Journal of Std & Aids,2014,26(3):145.

[4] Dreiseitl S,Ohno-machado L. Logistic regression and artificial neural network classification models: a methodology review. Journal of Biomedical Informatics,2002,35(5):352-9.

[5] 陈广,陈景武. logistic 回归分析的判别预测功能及其应用. 数理医药学杂志,2007,20(3):280-1.

[6] Kl Ppel B. Neural Networks as a New Method for EEG Analysis. Neuropsychobiology,1994,29(1):33-8.

[7] Breiman L. Random Forests. Machine Learning,2001,45(1):5-32.

[8] Ukil A. Support Vector Machine. Computer Science,2002,1(4):1-28.

[9] Efron B. Bootstrap Methods: Another Look at the Jackknife. 1979,7(1):1-26.

[10] Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers´D. Stata Journal,2010,10(3):339-58.

[11] Davis A,Best J,Luo J,et al. Risk behaviours,HIV/STI testing and HIV/STI prevalence between men who have sex with men and men who have sex with both men and women in China. Cambridge University Press,2015.

[12] Cortez P,Embrechts MJ. Using sensitivity analysis and visualization techniques to open black box data mining models. Elsevier Science Inc.,2013.

[13] Bengio Y. Practical Recommendations for Gradient-Based Training of Deep Architectures. Springer Berlin Heidelberg,2012.

[14] Olden JD,Jackson DA. Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling,2002,154(1):135-50.

(责任编辑:郭海强)