

基于贝叶斯分析的阶段性自主体育锻炼行为预测

曹 健¹, 孙龙龙²

(1.安徽理工大学,安徽 淮南 232001;2.宣城市第三中学,安徽 宣城 242000)

摘要:为了提高阶段性自主体育锻炼行为分析和判断能力,提出基于贝叶斯分析的阶段性自主体育锻炼行为预测方法。构建阶段性自主体育锻炼行为预测的统计时间序列分析模型,采用大数据特征检测方法进行体育锻炼行为大数据挖掘和特征提取,基于贝叶斯分析预测思想进行行为统计特征序列的有序聚类,结合模糊C均值聚类分析方法进行体育锻炼行为预测过程中的信息聚类 and 属性归并,提取统计时间序列的关联规则特征量,在加权马尔可夫链中实现对阶段性自主体育锻炼行为量的准确预测。仿真结果表明,采用该方法进行阶段性自主体育锻炼行为预测的准确性较高,提高了自主体育锻炼行为的量化分析能力。

关键词:贝叶斯分析;阶段性;自主体育锻炼;行为预测

中图分类号:G806 **文献标志码:**A **文章编号:**1673-1891(2020)04-0068-04

Prediction of Phased Autonomous Physical Exercise Behavior Based on Bayesian Analysis

CAO Jian¹, SUN Longlong²

(1.Anhui University Of Science & Technology,Huainan, Anhui 232001;

2.Xuancheng No.3 Middle School,Xuancheng, Anhui 242000,China)

Abstract: In order to improve the ability of analyzing and evaluating the behavior of phased independent physical exercise, this paper puts forward a prediction method based on Bayesian analysis. It constructs the statistical time sequence analysis model of phased independent physical exercise behavior prediction with the big data feature detection method to explore and extract the big data of physical exercise behavior. Based on the Bayesian analysis and prediction idea, orderly clustering of the behavior statistical feature sequence is carried out, and the information clustering and categorizing in the process of physical exercise behavior prediction are combined with fuzzy C-means clustering analysis method in this paper. Characteristic quantity of associated rule for the statistical time sequence is extracted, and the phased independent physical exercise behavior is accurately predicted in the weighted Markov chain. The simulation results show that the accuracy of this method is high, and the ability of quantitative analysis of autonomous physical exercise has been improved.

Keywords: Bayesian analysis; phased; autonomous physical exercise; behavior prediction

0 引言

随着我国经济水平的不断提升,人们的生产与生活方式已经发生了翻天覆地的变化,但是健康却已经亮起了红灯。据相关调查结果显示,我国一部分人群已经处于亚健康状态,尤其是青少年,由于手机、电脑以及其他上网设备的普及,越来越多青少年将更多的时间放在上网聊天以及打游戏方面,极大压缩了体育锻炼的时间,因此全民健康以及全面健

身问题已经成为了社会普遍关注的热点话题^[1]。

人们自觉参加体育锻炼受多种因素的影响,如个人的兴趣爱好、生活习惯、闲暇时间、硬件条件、氛围等^[2]。由于人的行为改变是一个复杂、渐进、连续的过程,进行一种新行为的个体将有序地进行,因此体育锻炼具有阶段性,也是有规律可循的。当前已有研究大多数都是根据问卷调查分析体育锻炼的态度状况及其与锻炼行为的关系以及体育锻炼行为动机冲突解释等,由于这些方法主观性较

收稿日期:2020-01-28

基金项目:安徽理工大学青年科学研究基金(文科类)(QN201646)。

作者简介:曹健(1989—),男,安徽淮南人,讲师,硕士,研究方向:高校体育教学、运动训练等。

强,因此难以达到理想的应用效果。

为了提升自主体育锻炼行为的量化分析能力,本文提出基于贝叶斯分析的阶段性自主体育锻炼行为预测方法。该方法研究体育锻炼行为所表现出的阶段性动力特征,在此基础上设计合理有效干预措施,促使体育锻炼由低级阶段向高级阶段有效转变,以达到全面提升国民身体素质的目的。

1 统计数据分析 and 特征提取

1.1 阶段性自主体育锻炼行为统计时间序列

阶段性自主体育锻炼行为过程的不确定性较强,为了实现对阶段性自主体育锻炼行为的准确预测,需要首先构建阶段性自主体育锻炼行为预测模型,提取阶段性自主体育锻炼行为的统计特征量,对阶段性自主体育锻炼行为信息进行特征提取,构建阶段性自主体育锻炼行为大数据分布结构模型^[3],得到阶段性自主体育锻炼行为的分布结构模型,特征状态方程描述为式(1)。

$$A(x) = AJ(x)a(x) + B(1 - b(x)) \quad (1)$$

令 $A = \{a_1, a_2, \dots, a_n\}$ 为阶段性自主体育锻炼行为的特征集, $B = \{b_1, b_2, \dots, b_m\}$ 为阶段性自主体育锻炼行为分布属性类别集, a_i 的属性值为 $\{c_1, c_2, \dots, c_k\}$, 采用一个连通的无向图 $G = (V, E, W)$ 表示阶段性自主体育锻炼行为分布, 阶段性自主体育锻炼行为预测的统计特征量之间的关系满足 $V = \{v_1, v_2, \dots, v_N\}$, $W = \{\omega_1, \omega_2, \dots, \omega_m\}$ 的权值集合。根据阶段性自主体育锻炼行为的统计特征分布进行自适应学习, 根据阶段性自主体育锻炼行为的统计时间序列分布进行高维空间重组^[4], 得到阶段性自主体育锻炼行为的统计时间序列分布的大数据分布矩阵为式(2)。

$$X = [x(t_0), x(t_0 + \Delta t), \dots, x(t_0 + (K-1)\Delta t)]$$

$$= \begin{bmatrix} x(t_0) & x(t_0 + \Delta t) & \dots & x(t_0 + (K-1)\Delta t) \\ x(t_0 + J\Delta t) & x(t_0 + (J+1)\Delta t) & \dots & x(t_0 + (K-1)\Delta t + J\Delta t) \\ \vdots & \vdots & \ddots & \vdots \\ x(t_0 + (m-1)J\Delta t) & x(t_0 + (1+(m-1)J)\Delta t) & \dots & x(t_0 + (N-1)\Delta t) \end{bmatrix} \quad (2)$$

式中, $X(t)$ 表示阶段性自主体育锻炼行为的特征分布序列, J 是扰动窗口函数, m 表示阶段性自主体育锻炼行为分布统计时间序列分布的维数。结合关联规则挖掘方法进行阶段性自主体育锻炼行为样本融合处理^[4], 得到阶段性自主体育锻炼行为的演化特征分布集式(3)。

$$S_b = \sum_{i=1}^c p_i (\vec{m}_i - \vec{m}) (\vec{m}_i - \vec{m})^T \quad (3)$$

其中 $\vec{m} = \sum_{i=1}^c p_i \vec{m}_i$ 是阶段性自主体育锻炼行为的统计特征集, 阶段性自主体育锻炼行为的关联规则集用 a_i 表示, 则阶段性自主体育锻炼行为概率分布

的简化模型可以用式(4)进行描述。

$$\begin{cases} G_1 = b_{11}a_1 + b_{12}a_2 + \dots + b_{1n}a_n \\ G_2 = b_{21}a_1 + b_{22}a_2 + \dots + b_{2n}a_n \\ \vdots \\ G_n = b_{n1}a_1 + b_{n2}a_2 + \dots + b_{nn}a_n \end{cases} \quad (4)$$

其中, G_j 和 G_k 都具有较强的关联性, 表示阶段性自主体育锻炼行为预测的过程具有可靠性。

1.2 自主体育锻炼行为特征提取

采用线性规划方法, 进行阶段性自主体育锻炼行为预测过程中的自适应寻优, 建立加权的马尔科夫预测的核函数 $K(x_i, x_j)$, 加权马尔科夫预测核函数模型为式(5)。

$$\left. \begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{j=1}^l \alpha_j \\ \text{s.t.} \quad & \sum_{j=1}^l y_j \alpha_j = 0 \\ & 0 \leq \alpha_j \leq u(x_j)C \quad j=1, 2, \dots, l \end{aligned} \right\} \quad (5)$$

对阶段性自主体育锻炼行为的状态特征进行自适应聚类^[5], 采用模糊聚类分析的方法, 得到阶段性自主体育锻炼行为的可靠性分布概率密度为式(6)。

$$w_{ij} = \beta \times w(e_p k_q) \quad (6)$$

其中, β 为自适应加权系数, $w(e_p k_q)$ 表示阶段性自主体育锻炼行为的波动可靠性系数, 计算阶段性自主体育锻炼行为的多维分布集, 为式(7)。

$$Y_k = [Y_{k1}, Y_{k2}, \dots, Y_{kj}, \dots, Y_{kN}] \quad (k=1, 2, \dots, N) \quad (7)$$

其中, Y_{kj} 表示稳态特征量, N 为数据长度, 依据不同指标间的关联性进行阶段性自主体育锻炼行为预测, 得到阶段性自主体育锻炼行为准确预测的概率密度分布^[6], 得到阶段性自主体育锻炼行为可靠性预测的均值函数式(8)。

$$AV_m(r) = \frac{1}{N - (m-1)\tau} \sum_{i=1}^{N-(m-1)\tau} \ln R_m(r, i) \quad (8)$$

其中, $ENTR = - \sum_{l=l_{\min}}^{N-1} P(l) \cdot \ln P(l)$ 在马尔科夫链中, 得到阶段性自主体育锻炼行为预测的误差为 $\delta_{ij} = \|x_i - x_j\|$, 当 $\|x_i - x_j\| < \varepsilon$, 得式(9)。

$$(J_{x_i}^{(0)})^T = B_{x_i}^\dagger B_{x_i}^{(0)} \quad (9)$$

在满足优化的聚类准则下, 得到阶段性自主体育锻炼行为特征聚类中心函数为 $X_{\eta(n)}$ 和 X_n , 采用大数据特征检测方法进行阶段性自主体育锻炼行为预测的大数据挖掘和特征提取, 基于贝叶斯分析预测思想, 进行自主体育锻炼行为预测^[7]。

2 阶段性自主体育锻炼行为预测优化

2.1 阶段性自主体育锻炼行为特征融合

结合模糊C均值聚类分析方法, 进行阶段性自

主体体育锻炼行为预测过程中的信息聚类 and 属性归并,提取阶段性主体体育锻炼行为统计时间序列的关联规则特征量,得到阶段性主体体育锻炼行为的离散数据集 $x(t)$,引入加权马尔科夫预测思想^[8],得到阶段性主体体育锻炼行为预测的主成分分量为式(10)。

$$z(t)=x(t)+iy(t)=a(t)e^{i\theta(t)} \quad (10)$$

式中: $z(t)$ 表示阶段性主体体育锻炼行为预测的主成分分量, $x(t)$ 表示阶段性主体体育锻炼行为预测的模糊特征分量, $y(t)$ 表示预测输出序列,构建阶段性主体体育锻炼行为预测的判别模型函数为式(11)。

$$\hat{W} = \begin{cases} \text{sgn}(W)(|W| - \alpha Ts) & |W| \geq Ts \\ 0 & |W| < Ts \end{cases} \quad (11)$$

基于贝叶斯分析预测思想,进行阶段性主体体育锻炼行为统计特征序列的有序聚类,得到阶段性主体体育锻炼行为信息融合模型为式(12)。

$$\begin{aligned} \max F(X) &= (F_1(X), F_2(X), \dots, F_n(X)) \\ \text{s.t. } g_j(X) &\leq 0 \quad (j=1, 2, \dots, p) \\ h_k(X) &= 0 \quad (k=1, 2, \dots, p) \end{aligned} \quad (12)$$

构建多尺度融合预测的加权学习模型,进行阶段性主体体育锻炼行为的可靠性预测,得到阶段性主体体育锻炼行为预测的自适应迭代函数为式(13)。

$$\begin{aligned} x_i^{(k+1)} &= (1-\omega)x_i^{(k)} + \frac{\omega}{a_{ni}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right) \\ i &= 1, 2, \dots, n \\ k &= 1, 2, \dots, n \end{aligned} \quad (13)$$

由此构建阶段性主体体育锻炼行为的预测统计分布模型,采用加权学习方法^[9],进行自适应寻优,得到模糊学习矩阵式(14)。

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \quad (14)$$

根据阶段性主体体育锻炼行为分布进行多维特征重构,得到阶段性主体体育锻炼行为预测的模糊关联规则分布集合为 $S=\{(x_1, y_1, u(x_1)), \dots, (x_n, y_n, u(x_n))\}$,其中 $x_j \in R^n$, $u(x_j) \in \{-1, 1\}$, $\sigma \leq u(x_j) \leq 1$, σ 为等效的贝叶斯分析模型参数,基于贝叶斯分析预测思想,进行阶段性主体体育锻炼行为统计特征序列的有序聚类分析。

2.2 主体体育锻炼行为预测优化

采用大数据特征检测方法进行阶段性主体体育锻炼行为预测的大数据分析和特征提取,基于贝叶斯分析预测思想,进行阶段性主体体育锻炼行为统计特征序列的概率密度分析^[10],得到阶段性自主

体育锻炼行为预测的概率密度分布函数为式(15)

$$h(t) = H \sum_{m=1}^M \sum_{k=1}^{K(m)} \alpha_{mk} \delta(t - T_m - \tau_{mk}) \quad (15)$$

根据阶段性主体体育锻炼行为特征变化,进行阶段性主体体育锻炼行为的统计分析,采用分段样本检验分析方法,得到阶段性主体体育锻炼行为的统计分析模型为式(16)。

$$\frac{\partial u_i}{\partial p_i} = \frac{G h_i}{\sum_{j \neq i} h_j p_j + \sigma^2} \left(\frac{1}{1 + \gamma_i} - \beta_{c_i} \right) \quad (16)$$

提取阶段性主体体育锻炼行为的统计特征量,采用高阶累积特征分析方法,进行阶段性主体体育锻炼行为的变化趋势预测,得到阶段性主体体育锻炼行为预测的统计特征量为式(17)。

$$\eta_{comm} = \frac{k_1 \cdot l}{E_{comm}} \cdot (1 - p_{drop}) \quad (17)$$

其中, p_{drop} 为阶段性主体体育锻炼行为的最小量化分布集,采用模糊聚类方法,得到阶段性主体体育锻炼行为预测的稳态特征量 $CH(i \in C_i)$,基于模糊C均值聚类,进行阶段性主体体育锻炼行为的统计分析和特征提取,输出为式(18)。

$$f(x) = \text{sgn} \left\{ \sum_{j=1}^l \alpha_j^* y_j K(x, x_j) + b^* \right\}, x \in R^n \quad (18)$$

其中, $b^* = y_i - \sum_{j=1}^l y_j \alpha_j K(x_j, x_i)$, $i \in \{i | 0 < \alpha_i^* < u(x_i)C\}$ 。

根据以上方法结合关联规则挖掘方法建立阶段性主体体育锻炼行为的概率密度分析模型,提取阶段性主体体育锻炼行为的特征量,结合模糊特征检测方法,实现对阶段性主体体育锻炼行为的概率密度分析和优化预测。

3 仿真测试分析

为了测试本文方法在实现阶段性主体体育锻炼行为优化预测中的应用性能,进行仿真测试分析,采用Matlab进行阶段性主体体育锻炼行为预测仿真分析,得到行为预测的结构分布(图1)。

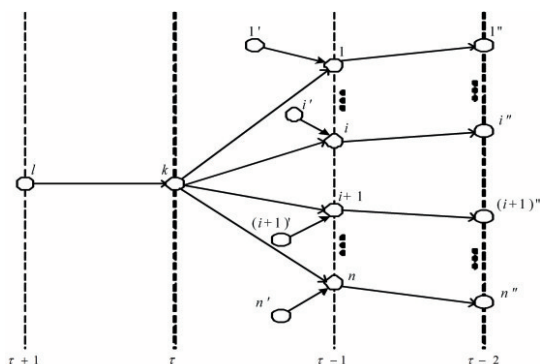


图1 行为预测的结构分布图

分析图1可知,阶段性自主体育锻炼行为预测的结构分为四层,每一层都能够对上一次结果进行细分,并对自主体育锻炼行为特征进行分类统计,获取行为的变化趋势,输出预测结果。

根据图1的行为预测的结构分布,进行阶段性自主体育锻炼行为预测,得到预测的收敛值(图2)。

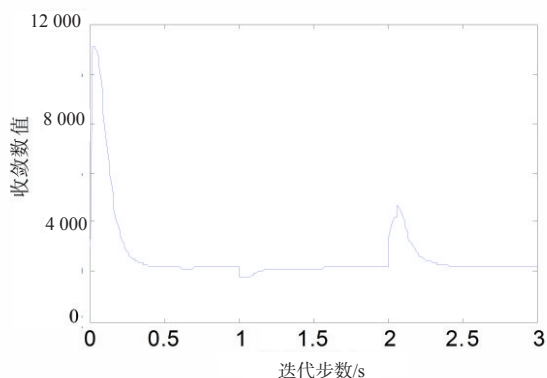


图2 阶段性自主体育锻炼行为预测

分析图2得知,利用本文方法进行阶段性自主体育锻炼行为预测的收敛性较好,主要原因在于该方法的收敛曲线拐点较少,因此采用该方法能够得到更为准确的预测结果。测试预测精度,得到对比结果(图3)。分析图3得知,本文方法的预测精度在100次迭代过程中始终优于传统方法,说明利用该

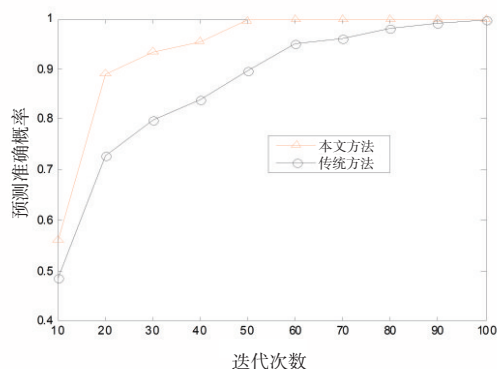


图3 阶段性自主体育锻炼行为预测准确概率

方法能够得到更为精准的阶段性自主体育锻炼行为预测结果。

4 结语

为了提升阶段性自主体育锻炼行为预测准确性,设计基于贝叶斯分析的阶段性自主体育锻炼行为预测方法。测试结果表明,本文方法的收敛性好,且对于阶段性自主体育锻炼行为的预测精度较高,因此将该方法应用至体育锻炼行为预测领域具有可靠性,能够进一步提升自主体育锻炼行为的量化分析能力,为我国国民体质健康的提升奠定坚实的基础。

参考文献:

- [1] 么广会,张龙,苏涛,等.大学生体育锻炼行为阶段性变化的实证研究[J].体育研究与教育,2016,31(2):73-77+85.
- [2] 李芊杉,王小莹.中学生运动动机与体育锻炼行为相关性研究——基于自我决定理论视角[J].体育研究与教育,2016,31(4):57-61.
- [3] 王硕.基于认知行为理论的大学生体育锻炼意向与锻炼行为的关系研究[J].体育世界:学术版,2017(12):85-87.
- [4] 张蓉.不同动机类型对大学生体育锻炼行为的预测[J].体育时空,2018(23):340-341.
- [5] 王新胜,卞震.基于贝叶斯模型的驾驶行为识别与预测[J].通信学报,2018,39(3):108-117.
- [6] 曾蔚,吴伊萍.一种基于卷积神经网络和贝叶斯网络的用户行为预测算法[J].安阳师范学院学报,2019(2):29-34.
- [7] 李柯,谭柱森,唐小艳.基于贝叶斯网络的应急物流风险预测与控制研究[J].物流科技,2017,40(2):98-101.
- [8] 余良芬.基于自我决定理论的女生课外体育锻炼动机特征及其与参与行为的关系研究[J].运动,2017(5):56-57.
- [9] 李铁波.基于校园大数据的学生行为特征分析与预测方法[J].重庆理工大学学报(自然科学版),33(7):201-206.
- [10] 孙丽娜,王应海,黄永红,等.基于贝叶斯推断LSSVM的枪管寿命建模与预测[J].火炮发射与控制学报,2018,39(4):93-97.