

协同训练算法在无创血糖检测中的应用

张迪¹, 陈真诚², 梁永波², 吴植强³, 朱健铭², 钟婷婷¹

1. 桂林电子科技大学电子工程与自动化学院, 广西 桂林 541004; 2. 桂林电子科技大学生命与环境科学学院, 广西 桂林 541004;
3. 广西壮族自治区医疗器械检测中心, 广西 南宁 530021

【摘要】目的:在无创血糖检测方法的研究中,因无创生理参数相比血糖真值更易于获取,病理数据库中未用血糖真值标记样本的数量远大于有标记的样本,若能将未标记样本应用于传统有监督血糖预测模型的训练中,将有效扩充训练样本集并提高模型的泛化能力。**方法:**在基于能量代谢守恒法的理论基础上,利用无创生理参数天然的多视图特性,将半监督学习算法应用于无创血糖的预测中,提出一种基于多视图协同训练与支持向量机技术的血糖预测算法。**结果:**经实验分析,在一定标记率下,基于协同训练的学习算法相比传统的有监督学习算法预测误差更小。说明未标记样本能够有效提升原始模型的泛化能力。**结论:**协同训练的引入,充分利用了规模较大的未标记样本,提高了模型泛化能力,并减少了血糖样本采集中标记样本的工作量,为今后无创血糖算法的研究提供了新思路。

【关键词】糖尿病;无创血糖检测;协同训练;支持向量机

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2018)11-1295-06

Application of co-training algorithm in noninvasive blood glucose detection

ZHANG Di¹, CHEN Zhencheng², LIANG Yongbo², WU Zhiqiang³, ZHU Jianming², ZHONG Tingting¹

1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; 2. School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, China; 3. Medical Devices Testing Center of Guangxi Zhuang Autonomous Region, Nanning 530021, China

Abstract: Objective In the study of noninvasive blood glucose detection, obtaining noninvasive physiological parameters is easier than obtaining true blood glucose. In the pathological database, the number of samples unlabeled with true blood glucose is much larger than that of labeled samples. This research aims to apply the unlabeled samples to the training of traditional supervised model for the prediction of blood glucose for effectively expanding the training set and improving the generalization ability of the model. **Methods** Based on the theory of conservation of energy metabolism and the natural multi-view characteristics of noninvasive physiological parameters, a semi-supervised learning algorithm was applied to the prediction of blood glucose. An algorithm based on multi-view co-training and support vector regression was proposed for the prediction of blood glucose. **Results** The experimental analysis showed that at a certain labeling rate, the prediction error of algorithm based on co-training is lower than that of traditional supervised learning algorithm, which indicated that unlabeled samples could effectively improve the generalization ability of the original model. **Conclusion** Co-training algorithm can fully utilize large-scale unlabeled samples, improve the generalization ability of model, and reduce the workload of labeling blood glucose samples, providing a new idea for the future research on noninvasive blood glucose algorithm.

Keywords: database; noninvasive blood glucose detection; co-training; support vector machine

【收稿日期】2018-09-20

【基金项目】国家自然科学基金重大科研仪器研制项目(61627807);广西自然科学基金(2017GXNSFGA198005);国家重点研发计划课题(2016YFC1305703);广西自然科学基金青年基金项目(2016GXNSFBA380145);广西自动检测技术与仪器重点实验室主任基金(YQ17118);广西信息科学实验中心一般项目(YB1513)

【作者简介】张迪,硕士研究生,研究方向:生物传感与智能仪器, E-mail: zd9258@gmail.com

【通信作者】吴植强,副主任技师,主要从事医疗器械检验检测和管理工作, E-mail: wzqnn@126.com

前言

糖尿病是一种常见的内分泌代谢疾病,其特点是葡萄糖脂肪代谢紊乱、血浆葡萄糖水平增高等。据国际糖尿病联盟(International Diabetes Federation, IDF)统计,2017年全球约4.25亿成人患糖尿病,预计到2045年,糖尿病患者可能达到6.29亿,且中国和印度是糖尿病患者最多的两个国家(占总数的48%)^[1]。糖尿病的治疗目前是尽量将人体血糖维持在正常的

浓度范围内,需要频繁进行血糖监测,因此血糖检测在糖尿病的防治中起到关键作用。

目前,常见的自我检测血糖仪可分为三类:第一类是微创血糖仪,包括激光式和针刺式等;第二类是连续式血糖监测仪(GCM),利用植入腹部皮下的感应器与组织间液中葡萄糖的反应,从而检测血糖趋势^[2-3];第三类是无创血糖检测仪,包括超声波法^[4]、光学方法^[5]、代谢热整合法、电阻抗谱等。其中,无创血糖检测具有无创伤、无痛苦、无感染等优势,给患者带来福音,因此对无创血糖检测方法的研究具有重要意义。

在无创血糖的研究过程中,因受试者对无创血糖检测接受程度远高于有创检测,大量获取无创血糖检测样本较为容易,而血糖真值相对较难。目前所用的病理数据库中,未用血糖真值标记的无创血糖参数样本数量远远高于有标记样本,若仅利用真值标记后的数据进行拟合,会面临数据浪费、样本量小等问题,若能同时将已标记和未标记样本参与到建模中,将极大地扩充用于建模的数据量,进一步提高血糖检测模型的泛化能力。

本文基于代谢热整合法,根据能量代谢守恒法的有关理论,在前人研究基础上,主要针对已标记数据相比未标记数据难以获取的问题,在支持向量回归机(SVR)的基础上,提出了基于协同训练的半监督无创血糖检测算法,能够同时利用血糖检测中的已标记样本和未标记样本,使得无创血糖检测系统仅需少量有创血糖值标记样本,即可保证整个样本集上模型检测的准确率。

1 算法原理

1.1 能量代谢守恒法无创血糖检测

日本学者Cho等^[6]提出了代谢热整合法,在该模型基础上,陈真诚等^[7-8]提出了能量代谢守恒法,根据人体局部产热推算代谢热交换速率,并由代谢热交换速率、血流速度、血氧等生理参数反推血糖值,该模型如下:

$$GLU = F_1((E, SPO_2, Bv, PF, \xi)) \quad (1)$$

式中 GLU 为预测血糖浓度, E 为局部热交换率, Bv 为血流速度, PF 为脉率, ξ 为修正项。其中热交换率 E 又可以表示为:

$$E = F_2(\Delta T, \Delta RH, \Delta R, \dots) \quad (2)$$

其中, ΔT 为手指表面与环境的对流换热量、 ΔRH 为手指皮肤蒸发换热量、 ΔR 为手指局部辐射散热, F_2 为热交换率算法,是 ΔT 、 ΔRH 、 ΔR 的多项式。整合后的血糖测定关系式为:

$$GLU = F(\Delta T, \Delta RH, \Delta R, E, SPO_2, Bv, PF, \xi) \quad (3)$$

目前,张杨等^[9]根据该关系式,完善了模型中的 F ,使血糖浓度预测值与真实值相关性达到了 88.5%。

1.2 半监督学习方法与协同训练

针对同时包含有未标记样本和已标记样本的数据集,半监督学习可以有效地利用大量未标记样本,对已标记样本构建模型起指导作用,在充分利用数据资源,减轻标记样本的工作量的同时,有效提高学习模型的性能^[10]。半监督学习主要包括自训练技术、基于图的半监督学习、S3VM、协同训练等技术^[11]。前三者基于单一学习策略,基于分歧的方法是利用多个学习器之间的分歧来提高对未标记样本分析能力的算法。半监督学习通常依赖于数据的聚类假设与流形假设,前者假设数据存在簇结构,要求同一个簇上具有相同的标签,通常聚类假设用于分类问题;后者假设数据在高维中表现为低维数据的流形嵌入,邻近的样本具有相差较小的输出。流形假设不要求输出具有二值性,适用于血糖值预测等回归问题。

协同训练(Co-training)是一种基于分歧的半监督学习方法,即构造两个具有分歧的弱学习器,在模型的每轮训练中,让每个弱学习器对未标记数据进行标注,选择其中置信度最高的样本加入另一个弱学习器的训练集中^[12-14]。通过这样相互提供训练样本的方式,最终使两个弱学习器的性能同时得到提升。协同训练的应用中,关键在于分歧的产生,常用方法是利用两个充分且条件独立的视图^[15],用两个视图分别构造基础弱学习器,“充分性”是指每个视图都足以完成弱学习器的训练;“条件独立性”是指在给定类别条件下,两个视图提供的信息相对独立。

2 实验分析与结果

2.1 研究对象与预处理

为了验证在少量标记样本情况下,无创血糖算法实际拟合效果,本文选用文献[9]使用的数据集 D ,即 300 例由微创血糖仪检测并标记真值的无创血糖检测样本。随机选取一定比例样本作为已标记样本 L ,其余样本隐去真实血糖值,作为未标记样本 U ,已标记样本占有所有样本的比例即数据的标记率(labeling rate)。

协同训练采用基础模型为基于 rbf 核的 SVR 回归器。在数据集 D 上随机选择 200 样本,利用遗传算法找出合适的惩罚系数 $C=9.25$ 与核参数 $\gamma=0.143$ 并训练得到 SVR 回归器^[16]。另 100 样本用该回归器

进行血糖值预测。经比较,100个测试样本预测血糖值与血糖真实值拟合相关性为0.850°,略低于文献中“DE-PLS”算法的拟合相关性0.885,但是单SVR回归器具有泛化能力强、时间开销小等重要优势,且区别于原文中“先分类,后拟合”的学习策略,利于协同训练算法的开展。

该数据集中各属性量级差别较大,最大的参数取值达到 10^9 ,为避免特征之间的量级差异给学习过程带来的不利影响,用于SVR回归器的数据首先需要对输入特征进行标准化处理,过程如图1所示。

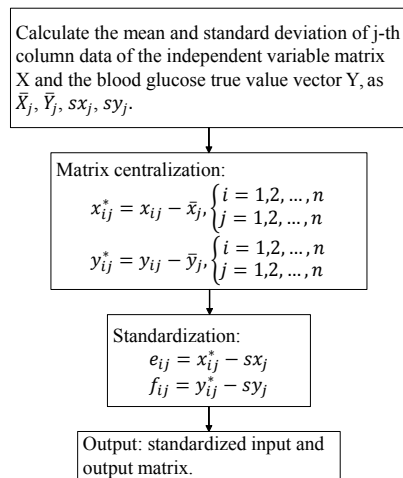
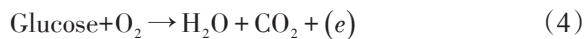


图1 数据标准化流程图

Fig.1 Flow chart of data standardization

2.2 基于多视图方法的协同训练SVR算法设计

最初的协同训练算法上,要求具有分歧的两个视图具有充分性和条件独立性,最好是数据集本身具有的天然视图分割。在上文所述基于能量守恒法检测无创血糖的原理中,式(1)表示了反推血糖值同时用到了能量代谢和耗氧量。根据如下基本反应式:



可以得知耗氧量和能量代谢理论上可以单独作为反推Glucose的充分条件,满足协同训练对两视图充分性的要求,其中 e 为产能量,反应式中各物质、能量比例是一定的。在传统的协同训练基础上有学者证明,条件独立性较弱的视图也能使协同学习器的性能得到提升^[17],且基于光电的心血管生理参数检测与其他环境、人体表面参数检测有着本质区别。因此将式(2)改写为:

$$\begin{cases} GLU = F(E, O, g) \\ E = \{HE, \Delta T, \Delta RH, \Delta^4 R, \dots\} \\ O = \{Hb, SPO_2, Bv, PF, \dots\} \end{cases} \quad (5)$$

其中 g 包含了该受试者性别、年龄等通用参数; E 包含了检测到的与能量代谢有关的参数,如传热系数

HE 、人体环境温度差 ΔT 等; O 包含了与耗氧量有关的生理参数,如血红蛋白 Hb 、血氧饱和度 SPO_2 等。利用视图 E 不用视图 O 在数据集上进行有监督学习得到弱学习器 h_1 ,利用视图 O 不用视图 E 在数据集上对 GLU 进行有监督学习得到弱学习器 h_2 ,即 $GLU = h_1(E, g) = h_2(O, g)$,利用未标记数据集, h_1 与 h_2 可以相互学习提高性能,基于SVR的多视图协同训练的血糖预测算法训练过程如表1所示。

该算法的关键是两个视图的选取。视图 $E = \{her, \Delta T, \Delta RH, \Delta^4 R, \dots\}$ 、 $O = \{Hb, SPO_2, Bv, PF, \dots\}$ 的选择是基于检测原理自然形成的分歧,各特征分量之间无显著相关,符合多视图学习对视图条件独立性的要求。另外对 E 、 O 和其他相关参数做基于偏最小二乘回归的各变量投影重要性(Variable Importance for Projection, VIP)分析,结果如图2所示,每个变量的VIP值一定程度上反映了自变量对血糖真值的解释作用^[18]。由图2可以看出,视图 E 主要参数 ΔR 、 $R_s^4 - R_E^4$ 、 HE ,以及视图 O 主要参数 SPO_2 、 Bv 的VIP值均高于1,被认为对预测无创血糖值线性贡献度较大,即视图 E 与视图 O 同时包含足以产生最优学习器的信息。

该算法的另一个关键在于,利用 h'_1 、 h'_2 对未标记数据进行标记时,如何选取置信度最高的标记样本加入另一个弱学习器的训练集中。根据半监督学习的流形假设,置信度高的样本应与已标记样本集里与其标记值较近的样本具有较高的相似程度。因此,首先计算已标记样本用 h'_1 、 h'_2 预测的均方误差 mse_1 、 mse_2 ,再将缓冲池中每个未标记样本 x_i 单独加入训练集后训练得到 $h_1^{(i)}$ 、 $h_2^{(i)}$, $(i = 1, 2, \dots, |U|)$,计算已标记样本用 $h_1^{(i)}$ 、 $h_2^{(i)}$ 预测的均方误差 $mse_1^{(i)}$ 、 $mse_2^{(i)}$, $(i = 1, 2, \dots, |U|)$,将每个 $mse_1^{(i)}$ 、 $mse_2^{(i)}$ 与 mse_1 、 mse_2 进行比较,因为置信度较高的样本的加入会使学习器提高泛化能力,降低均方误差,即 $\delta_j^{(i)} = mse_j - mse_j^{(i)} > 0$, $(i = 1, 2, \dots, |U|; j = 1, 2)$ 。因此每轮迭代选择使 $\delta_1^{(i)}$ 、 $\delta_2^{(i)}$ 为正且最大的样本 $x_1^{(i)}$ 、 $x_2^{(i)}$ 即为置信度最高样本。

3 实验结果分析

为了评价算法优劣和验证在不同标记率下该算法建模的回归效果,依次按照3%、6%、10%、15%、20%、30%、40%、50%标记率将 D 随机划分为 L 和 U ,在不同标记率下对 D 进行半监督SVR训练,记录模型对 U 的预测结果,并将各标记率下传统SVR方法仅对未标记数据训练结果作为参考,这里采用均

表 1 基于SVR的多视图协同训练的血糖预测算法

Tab.1 Multi-Viewed Co-training style blood glucose regression algorithm based on SVR

输入: 已标记数据 L , 样本数记为 $|L|$; 未标记数据 U , 样本数记为 $|U|$; 最大迭代次数 T ; 缓冲池大小 s

过程: 初始化两个训练集 $L_1=L_2 \leftarrow L$, 两个弱学习器 $h_1 \leftarrow \text{SVR}(L_1(E, g))$, $h_2 \leftarrow \text{SVR}(L_2(O, g))$

从 U 中随机选 s 个样本作为缓冲池 U_s , $U \leftarrow U \setminus U_s$

迭代 T 次:

分别计算 L_1 、 L_2 用模型 h_1 、 h_2 预测的均方误差 mse_1 、 mse_2

用 h_1 、 h_2 分别对 $U_s(E, g)$ 、 $U_s(O, g)$ 进行预测, 得到预测标记 u_1 , u_2

循环 $|U_s|$ 次: ($i = 1, 2, \dots, |U_s|$)

$h'_1 \leftarrow \text{SVR}(\{L_1 \cup \{u_1^{(i)}, U_s^{(i)}\}(E, g)\})$, $h'_2 \leftarrow \text{SVR}(\{L_2 \cup \{u_2^{(i)}, U_s^{(i)}\}(O, g)\})$

分别计算 L_1 、 L_2 用模型 h'_1 、 h'_2 预测均方误差 $mse_1^{(i)}$ 、 $mse_2^{(i)}$

$\delta_1^{(i)} = mse_1 - mse_1^{(i)}$, $\delta_2^{(i)} = mse_2 - mse_2^{(i)}$

对 $j = 1, 2$, 若 $\max(\delta_j^{(i)}) > 0$, $L_{3-j} \leftarrow L_{3-j} \cup \text{argmax}_{u \in U_s}(\delta_j^{(i)})$

若上一步中 L_1 、 L_2 均未发生改变, 则跳出迭代

更新两个弱学习器 $h_1 \leftarrow \text{SVR}(L_1(E, g))$, $h_2 \leftarrow \text{SVR}(L_2(O, g))$

从 U 中随机选择样本加入缓冲池 U_s 使 U_s 维持大小

输出模型: $h(x) = 0.5 * (h_1(x(E, g)) + h_2(x(O, g)))$

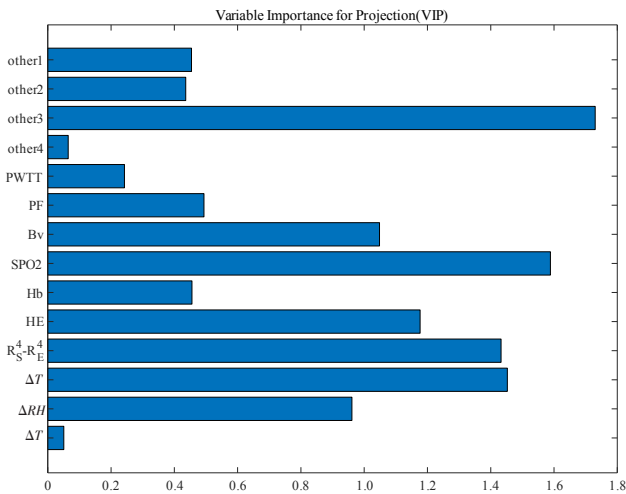


图 2 主要血糖参数的VIP分析图

Fig.2 Variable importance for projection of major blood glucose parameters

方误差与相关性作为度量指标,用以评价半监督学习对模型性能是否具有提升作用。实验结果如表 2 所示。

从表 2 中可以看出,标记率在 10%~40% 区间时,协同训练可以使原 SVR 算法误差下降,提高了模型泛化能力;标记率在 10% 以下和 40% 以上的情况下,协同训练未能有效提升原算法。

表 2 不同标记率下实验结果(参数: $s=50$; $T=200$)

Tab.2 Test results at different labeling rates (parameter settings: $s=50$; $T=200$)

Labeling rate/%	SVR		SVR based on co-training	
	Mean square error	Correlation	Mean square error	Correlation
3	9.50	0.14	18.95	0.11
6	7.60	0.32	8.07	0.23
10	4.87	0.63	4.70	0.68
15	5.01	0.68	3.19	0.76
20	4.59	0.74	2.49	0.81
30	2.89	0.81	2.51	0.82
40	2.33	0.83	2.32	0.83
50	2.18	0.84	2.30	0.83
66*	2.10	0.85	-	-

*Supervised learning, described in section 2.1 ($|L|=200$, $|U|=100$);
SVR: Support vector regression

为了探究协同训练的迭代是否确实能提高两个弱学习器的泛化能力,以 10% 标记率的实验为例,每隔 10 轮记录 h_1 、 h_2 在当前各自训练集 L_1 、 L_2 上的训练均方误差 mse_1 、 mse_2 , 作图显示其趋势如图 3 所

示,可以看出通过多轮迭代,两个弱分学习器在各自训练集上均能有效降低其训练均方误差。

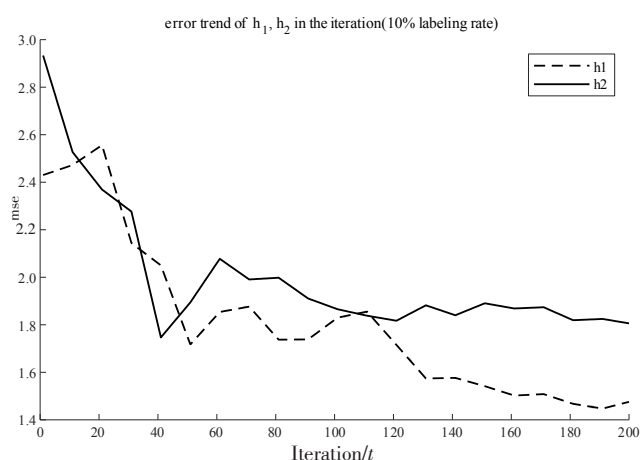


图3 10%标记率实验中 h_1 、 h_2 性能变化趋势图

Fig.3 Error trends of h_1 and h_2 at 10% labeling rate

为了观察每轮迭代产生的决策模型在未标记集 U 上性能的具体表现,将各标记率下训练迭代的过程可视化,每次迭代后额外计算目前 $h'(x) = 0.5 \times (h_1'(x(E, g)) + h_2'(x(O, g)))$ 在 U 上预测均方误差 mse ,该结果一定程度上反映了当前决策模型泛化能力,结果如图4所示,横坐标表示迭代次数,纵坐标表示当前决策模型预测误差。

4 结论

标记率小于10%时,半监督回归未能有效提高模型泛化能力,原因应为随机选取的用于初始建模的有标记数据点过少(9个、18个),随机性过强,该样本集的标签未能起到有效指导作用,基于rbf核的SVR回归在样本量极少的情况下容易使数据过于集中,在后续多轮迭代中也无法提升模型的泛化能力;随着标记率的增大,半监督回归对模型性能改善能力

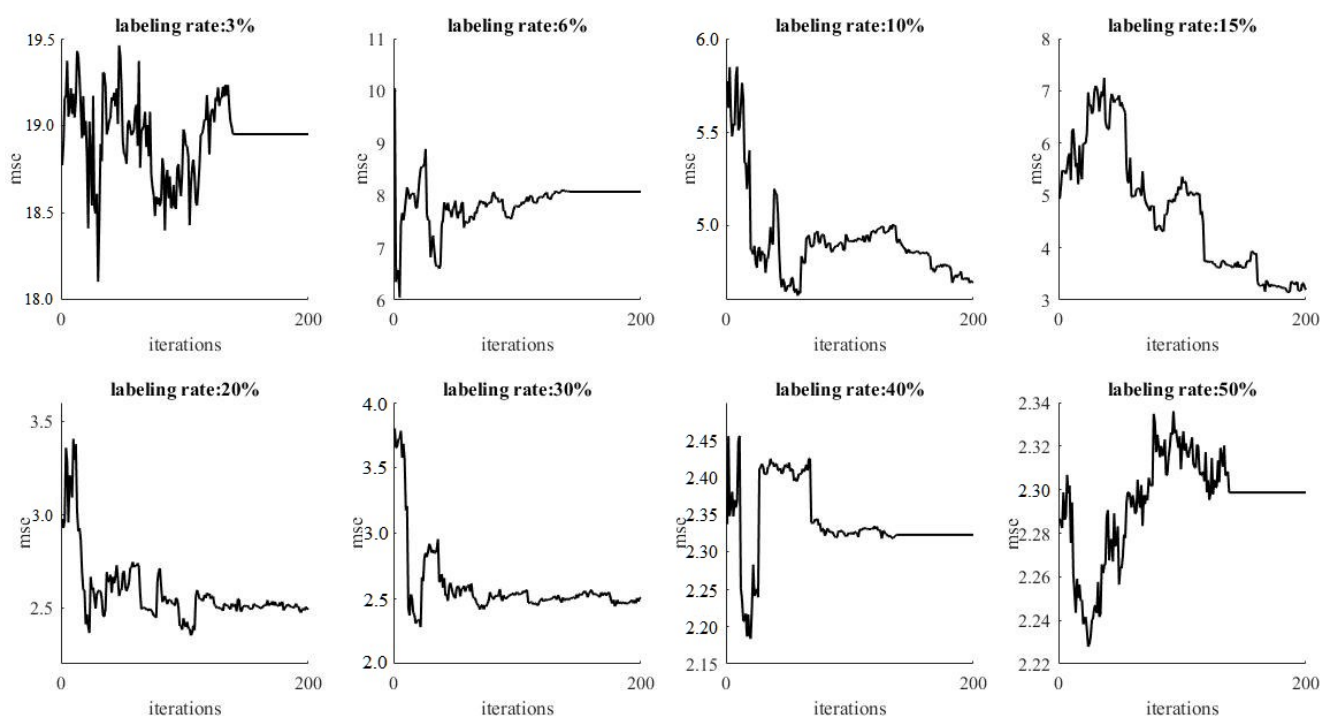


图4 每轮迭代对模型性能的影响

Fig.4 Effects of each iteration on model performance

逐渐体现,验证了半监督学习对无创血糖数据进行的流形假设,意味着相近的血糖值在特征空间中具有相似的表现;标记率大于等于40%时,半监督回归模型与单纯的SVR回归模型性能接近,多次迭代未能使模型泛化能力得到显著提升,即未标记样本的引入对现有模型提升意义不大。可以认为,在总样本量为300、标记率约为10%~20%的条件下,半监督回归对模型泛化能力提高有显著作用。

本文利用了无创血糖参数的多视图特性,提出了基于协同训练的SVR无创血糖算法,仅需20%~30%的有标记血糖样本即可使模型基本达到有监督训练的效果,并验证了血糖参数中存在的流形假设,降低了今后无创血糖算法研究对有标记血糖样本的需求,有效利用了易于获取的未标记血糖样本,为无创血糖算法研究提供了新思路。

【参考文献】

- [1] International Diabetes Federation. IDF diabetes atlas[EB/OL]. 8th ed. Brussels, Belgium: International Diabetes Federation, 2017. <http://www.diabetesatlas.org>.
- [2] 李强, 郭琳. 动态血糖监测的准确性与临床应用[J]. 中国实用内科杂志, 2016, 36(10): 833-837.
LI Q, GUO L. The accuracy of blood glucose monitoring and its clinical application [J]. Chinese Journal of Practical Internal Medicine, 2016, 36(10): 833-837.
- [3] 余丽玲, 陈婷, 金浩宇, 等. 基于支持向量机和自回归积分滑动平均模型组合的血糖值预测[J]. 中国医学物理学杂志, 2016, 33(4): 381-384.
YU L L, CHEN T, JIN H Y, et al. Prediction of blood glucose level based on model combining support vector machine and autoregressive integrated moving average[J]. Chinese Journal of Medical Physics, 2016, 33(4): 381-384.
- [4] CHOWDHURY M K, SRIVASTAVA A, SHARMA N, et al. Noninvasive blood glucose measurement utilizing a newly designed system based on modulated ultrasound and infrared light [J]. Int J Diabetes Dev Ctries, 2016, 36(4): 439-448.
- [5] 余振芳. 光学无创血糖检测技术研究[D]. 成都: 电子科技大学, 2017: 13-16.
YU Z F. Optical non-invasive blood glucose detection technology research [D]. Chengdu: University of Electronic Science and Technology of China, 2017: 13-16.
- [6] CHO O K, KIM Y O, MITSUMAKI H, et al. Noninvasive measurement of glucose by metabolic heat conformation method [J]. Clin Chem, 2004, 50(10): 1894-1898.
- [7] 陈真诚, 龚琼, 王弟亚, 等. 基于能量守恒的无创血糖检测传感器的设计[J]. 传感技术学报, 2008, 21(5): 724-728.
CHEN Z C, GONG Q, WANG D Y, et al. Design of non-invasive blood sugar measuring sensor based on energy metabolic conservation[J]. Chinese Journal of Sensors and Actuators, 2008, 21(5): 724-728.
- [8] 朱健铭, 陈真诚. 能量代谢守恒法无创血糖检测算法研究[J]. 传感技术学报, 2013, 26(7): 917-921.
ZHU J M, CHEN Z C. Research on the algorithm of non-invasive blood glucose measurement based on the method of conservation of energy metabolism [J]. Chinese Journal of Sensors and Actuators, 2013, 26(7): 917-921.
- [9] 张杨, 陈真诚, 朱健铭, 等. 一种改进的能量代谢守恒法无创检测血糖算法[J]. 传感技术学报, 2016, 29(6): 808-812.
ZHANG Y, CHEN Z C, ZHU J M, et al. An improved method for noninvasive detection of blood glucose of conservation of energy metabolism[J]. Chinese Journal of Sensors and Actuators, 2016, 29(6): 808-812.
- [10] CHAPELLE O, ZIEN A. Semi-supervised learning (adaptive computation and machine learning)[M]. The MIT Press, 2006.
- [11] ZHU X, GOLDBERG A. Introduction to semi-supervised learning[J]. Appl Stat, 2017, 172(2): 1826-1831.
- [12] 马蕾, 汪西莉. 基于支持向量机协同训练的半监督回归[J]. 计算机工程与应用, 2011, 47(3): 177-180.
MA L, WANG X L. Semi-supervised regression based on support vector machine cooperative training [J]. Computer Engineering and Applications, 2011, 47(3): 177-180.
- [13] 郭翔宇, 王魏. 一种改进的协同训练算法: compatible co-training[J]. 南京大学学报(自然科学), 2016, 52(4): 662-671.
GUO X Y, WANG W. An improved co-training style algorithm: compatible co-training[J]. Journal of Nanjing University (Natural Science), 2016, 52(4): 662-671.
- [14] LUO W P, LI H Q, SHI N. Semi-supervised least squares support vector machine algorithm: application to offshore oil reservoir[J]. J Appl Geophy, 2016, 13(2): 406-415.
- [15] 唐静静, 田英杰. 多视角学习综述[J]. 数学建模及其应用, 2017, 6(3): 1-15.
TANG J J, TIAN Y J. A survey on multi-view learning [J]. Mathematical Modeling and its Application, 2017, 6(3): 1-15.
- [16] NIE F, LI J, LI X. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification [C]//International Joint Conference on Artificial Intelligence. AAAI Press, 2016: 1881-1887.
- [17] 周志华. 基于分歧的半监督学习[J]. 自动化学报, 2013, 39(11): 1871-1878.
ZHOU Z H. Disagreement-based semi-supervised learning[J]. Acta Automatica Sinica, 2013, 39(11): 1871-1878.
- [18] 张政, 冯国双. 变量投影重要性分析在自变量筛选中的应用[J]. 现代预防医学, 2012, 39(22): 5813-5815.
ZHANG Z, FENG G S. Application of variable importance for projection in the variables selection[J]. Modern Preventive Medicine, 2012, 39(22): 5813-5815.

(编辑: 薛泽玲)