

基于无创血糖的糖尿病评估

陈真诚¹, 杨薛冰¹, 邹春林², 严波文¹, 朱健铭³, 梁永波³

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. 广西医科大学转化医学研究中心, 广西 南宁 530021; 3. 桂林电子科技大学生命与环境科学学院, 广西 桂林 541004

【摘要】为了提高糖尿病前期的检出率,在糖耐量受损(IGT)常规诊断方法的基础上,增加糖化血红蛋白作为糖尿病筛查的因素,构建一个IGT检测模型。采集受试者的身高、体质量、腹围、血压、皮脂厚度、空腹血糖和糖化血红蛋白作为模型的特征输入,用K-近邻算法和神经网络对其分类,模型输出包括血糖值正常、IGT和糖尿病。结果显示增加糖化血红蛋白作为分类特征后,神经网络和K-近邻算法的分类准确率分别为88.89%和93.09%,明显高于传统方法的分类准确率(83.33%和78.38%)。本研究提出的IGT检测模型对糖尿病的临床诊断有重要意义。

【关键词】糖尿病;糖化血红蛋白;糖耐量受损;神经网络;K-近邻算法

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2020)10-1330-05

Diabetes evaluation based on noninvasive blood glucose

CHEN Zhencheng¹, YANG Xuebing¹, ZOU Chunlin², YAN Bowen¹, ZHU Jianming³, LIANG Yongbo³

1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; 2. Transforming Medical Research Center, Guangxi Medical University, Nanning 530021, China; 3. School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, China

Abstract: On the basis of routine methods for diagnosing impaired glucose tolerance (IGT), an IGT detection model is constructed by adding glycosylated hemoglobin as the factor of diabetes screening, thereby improving the detection rate of pre-diabetes. The height, body weight, abdominal circumference, blood pressure, thickness of sebum, fasting blood glucose and glycosylated hemoglobin of subjects were collected as the feature inputs of the model, and then the subjects were classified by K-nearest neighbor algorithm and neural network. The output of the model included normal blood glucose, IGT and diabetes. The results show that after adding glycosylated hemoglobin as the factor of diabetes screening, the classification accuracies of neural network and K-nearest neighbor algorithm are 88.89% and 93.09%, respectively, which are significantly higher than 83.33% and 78.38% of traditional methods. The proposed IGT detection model is of great significance for the clinical diagnosis of diabetes.

Keywords: diabetes; glycosylated hemoglobin; impaired glucose tolerance; neural network; K-nearest neighbor algorithm

前言

糖尿病在日常生活中很常见,是一种高血糖非急性代谢性疾病^[1]。目前尚未有治愈糖尿病的方法,现阶段只能尽最大可能防止更多的人患糖尿病。我国糖尿病人数是世界之最,这将会对我国造成严重的影响^[2]。很多糖尿病前期患者的初期患病症状不

易被察觉,因此会错过预防的黄金期,最后发展成糖尿病患者^[3]。如果能及时诊断糖尿病前期患者,每年6%~10%的糖耐量受损者能转回正常人,不会发展成为糖尿病患者^[4]。糖耐量是指机体对血糖浓度的调节能力^[5],国内外研究表明,与糖耐量正常者相比,糖耐量受损者有更多的健康安全隐患,糖耐量受损产妇生出畸形儿的概率更大,且糖尿病前期人群更容易患心血管疾病^[6-7]。糖耐量受损患者的及时诊断和预防是阻断糖尿病人群扩大的一个途径^[8-9]。口服糖耐量试验(Oral Glucose Tolerance Test, OGTT)是临床上常用的糖尿病检查方法。OGTT需要多次在静脉采集血液,会对患者造成一定的心理和身体压力;现在国内外也有多种简单易行的非损伤糖尿病筛查风险评估模型^[10]和糖尿病简易检测方法^[11]。本研究

【收稿日期】2020-04-11

【基金项目】国家自然科学基金重大仪器研制项目(61627807);国家自然科学基金(1873913);广西自然科学基金(2018GXNSF-DA281044);广西创新研究团队项目(2017GXNSF-GA198005)

【作者简介】陈真诚,教授,研究方向:生物传感与智能仪器, E-mail: chenzhcheng@163.com

在糖尿病高危因素的基础上结合糖化血红蛋白,利用K-近邻(K-Nearest Neighbor, KNN)算法和神经网络对糖尿病进行分类,为日后辅助临床筛选糖尿病患者打下基础。

1 算法模型

1.1 神经网络

神经网络又称人工神经网络,是多学科交叉的一个领域。神经网络是通过许多结构类似的神经元连接的,其中最基本的成分是神经元模型,其结构简单易懂,而它组成的神经网络结构却是非常复杂的。单个神经元模型的结构如图1所示^[12-13]。

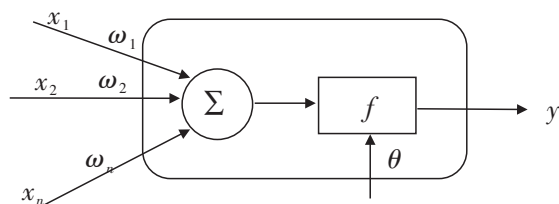


图1 单个神经元结构图
Fig.1 Single neuron structure diagram

图1中, x_1, x_2, \cdots, x_n 是前 n 个神经元传递来的输入向量,这些信号经过传递,产生神经元的输出向量,即 y ; $\omega_1, \omega_2, \cdots, \omega_n$ 是各个神经元之间的权系数; θ 为神经元的外部偏置,如果神经元的电位超过了一个外部偏置,它就会被激活,向其他神经元传递信息; f 是激活函数,理想中的激活函数是阶跃函数。阶跃函数的表达式为:

$$\text{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

(1)

不过由于其特性不光滑,不连续。所以实际常用sigmoid函数。sigmoid函数的表达式为:

$$\text{sigmoid}(x) = \frac{1}{1 + \exp(-t)}$$

(2)

从图1得到神经元的数学模型表达式为:

$$y = f\left(\sum_{i=1}^n x_i \omega_i + \theta\right)$$

(3)

神经网络一般包括一个输入层、一个或者多个隐含层和一个输出层,而隐含层是整个神经网络的核心。

1.2 KNN算法

KNN算法是由Hart和Cover等在1968年提出的^[14]。KNN算法意味着每个样本可以用最接近它的 K 个样本来表示,并可通过测量不同特征值之间的距离来进行分类。由于KNN算法的规则比较简单,且运行效率也高,所以KNN算法在分类聚类方面有着广泛的应用。图2为是KNN算法流程图。

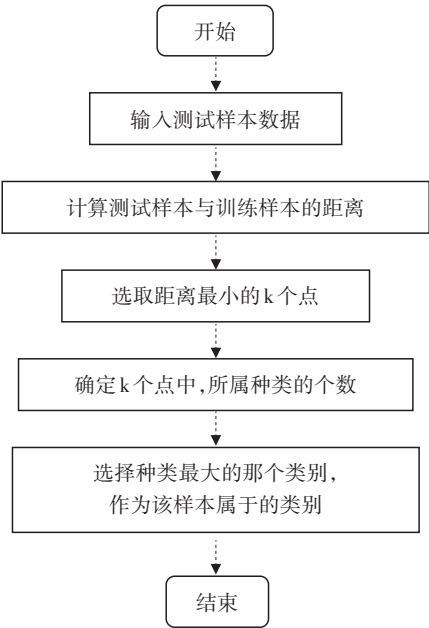


图2 KNN算法流程图
Fig.2 K-nearest neighbor (KNN) algorithm flowchart

2 实验

2.1 数据来源及采集

公开召集54名正常受试者、11名糖耐量受损者和41名糖尿病患者,年龄20~61岁;因为根据医学规定,糖耐量受损者和糖尿病患者在医学上是不能再进行糖耐量实验的,所以糖耐量受损者和糖尿病患者的数据均来自广西医科大学。广西医科大学采集生理参数的实验器材、实验要求与我们采集正常受试者的实验器材、实验要求均保持一致。

54名正常受试者的数据采集如下,在实验开始前一天从晚上10点起可以饮少量的水,禁止饮食。实验当天早晨6点到达实验地点,休息30 min,使受试者处于平静状态。6点30分实验开始,用家用人体电子秤(永康市乐天电子有限公司,型号LOT-28a)测量每个人的体质量,并做数据记录,体质量保留两位小数;采用皮脂厚度计(常熟市新枫仪表有限公司,型号PZJ_01)测量肚脐部右1 cm处的腹部皮脂厚度,共测3次,取其平均;用市尺(YingShi, 型号82223)测量肚脐下1 cm左右处,测量腹围3次,并做好数据记录。常规生理参数采集结束后,6点40分开始进行OGTT,采集血糖数据,并用无创GM-A血糖仪(桂林永成医疗科技有限公司,型号11171L0003)测量0~3 h中7个时间点血糖值,并做好记录。OGTT的试验值反映机体对超生理量的血糖的调节能力。

第二天组织正常受试者在广西壮族自治区桂林市医学院附属医院进行糖化血红蛋白的测量。血红蛋白和血糖相结合形成糖化血红蛋白,反映机体2月

~3月前的血糖水平。

全部受试者(正常受试者、糖耐量受损者、糖尿病患者)均知晓并同意参与本研究。排除标准包括恶性肿瘤患者、免疫性糖尿病患者、肾功能障碍患者、凝血系统功能不全患者以及患有糖尿病复发病的患者。

2.2 数据汇总及处理

无创血糖仪每个通道对应一个受试者的信息,从无创血糖仪里面导出每个通道对应的数据,并统计每一个受试者的身高、体质量、腹围和皮脂厚度。

国外有很多糖尿病风险评分工具是通过问卷或者计算公式得分等级等方式来判断普通群体是否健康,使用规则简单易学,并且没有创伤性。同时,也有很多国家根据国民的身体情况建立了本国的糖尿病风险评分工具,如芬兰糖尿病风险评分法等^[15]。我国虽然也有糖尿病风险评分模型^[16]以及一些改进之后的评分模型等^[17],但目前尚未建立统一的糖尿病评分工具。在常规诊断方法的基础上,本研究加入糖化血红蛋白作为判断标准,根据空腹血糖、体质量、身高、血压、腹围、皮脂厚度及糖化血红蛋白来进行人群划分,分为正常、糖尿病前期、糖尿病^[18-19]。具体实验数据如表1所示。

表1 不同分组的基本信息
Tab.1 Basic information of different groups

指标	正常对照组	糖尿病前期组	糖尿病组
年龄/年	28.20±9.54	52.67±13.56	48.39±11.19
性别(男/女)	27/27	2/9	22/19
身高/cm	167.22±8.00	162.33±6.71	167.32±7.04
体质量/kg	62.12±11.49	59.92±11.77	63.95±10.73
心率/次·min ⁻¹	72.46±8.43	76.08±4.52	78.45±6.40
舒张压/mmHg	69.95±7.98	77.75±4.92	81.63±7.10
收缩压/mmHg	116.11±12.41	120.83±8.05	123.12±11.20
空腹血糖/mmol·L ⁻¹	5.16±0.45	5.98±0.60	9.60±2.94
腹围/cm	79.29±7.74	78.29±9.54	80.96±7.65
皮脂厚度/mm	13.20±3.06	14.42±2.17	13.88±2.06
糖化血红蛋白/%	5.07±0.34	5.90±0.53	8.50±2.21

排除异常数据,如很明显的记录数据错误。特征参数的计量单位不同会导致训练过程特征参数的权重失去比较的意义,所以在训练特征参数前,为了使没有可比性的数据变得具有可比性,但同时保持相比较的两个数据之间的相对关系,需要对特征参

数进行归一化处理。

因为该实验无创测得血糖是手指指端的血糖,根据世界卫生组织(WHO)1999年的糖尿病诊断标准,指端血糖判定标准如表2所示。

表2 糖尿病毛细血管全血判定标准(mmol/L)
Tab.2 Criteria for determination of diabetes by capillary blood glucose (mmol/L)

分组	末梢血空腹血糖	糖负荷2 h末梢血糖
健康组	<5.6	<7.8
空腹血糖受损	≥5.6且≤6.0	<7.8
糖耐量低减	<6.1	≥7.8且≤11.0
糖尿病组	≥6.1	≥11.1

3 结果

对采集到的106组数据,按照3:2的比例划分,分别对应训练数据和测试数据,即64组为训练数据,42组为测试数据。在MATLAB 2019b上使用KNN算法进行分类研究,得到加入糖化血红蛋白的分类准确率和未加入糖化血红蛋白的分类准确率,运行的结果如图3和图4所示。加入糖化血红蛋白之后,KNN算法的准确率得到明显提高,可达93.09%;未加入糖化血红蛋白的分类准确率较低,仅为78.38%。

同时按照同样的流程,在MATLAB 2019b上使用神经网络对数据做分类。在神经网络中也同样分别输入8个特征参数和7个特征参数,分类结果见图5和图6。神经网络加入糖化血红蛋白之后,准确率有提高,达到88.89%;未加入糖化血红蛋白的分类准确率较低,仅为83.33%。可见加入糖化血红蛋白后,神经网络可以提高对糖尿病的分类准确率。

4 结论

把糖化血红蛋白作为特征输入之后,模型分类的准确率得到了提高。这种无创采集血糖的方式更加方便与快捷,并且降低了受试者的精神和身体压力。研究结果也表明利用KNN算法或神经网络搭建的模型对糖尿病分类的准确率较高,可以辅助临床诊断糖尿病。

本次采集的受试者人数有限,数据量不大,后期可以增加数据量,使模型更加完善。采集的皮脂厚度为腹部的皮脂厚度,以后可以考虑采用胳膊的皮脂厚度、大腿的皮脂厚度、腹部的皮脂厚度,综合计算人体的脂肪含量,将脂肪含量做为参数,使得模型更加精准。此外,实验还可以考虑增加检测指标,如心血管疾病、糖尿病家族史等,以进一步完善模型。

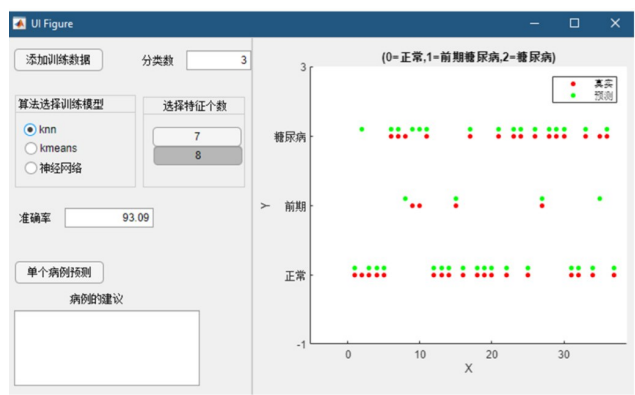


图3 KNN算法的分类准确率(8个输入参数)
Fig.3 Classification accuracy of KNN algorithm (8 input parameters)

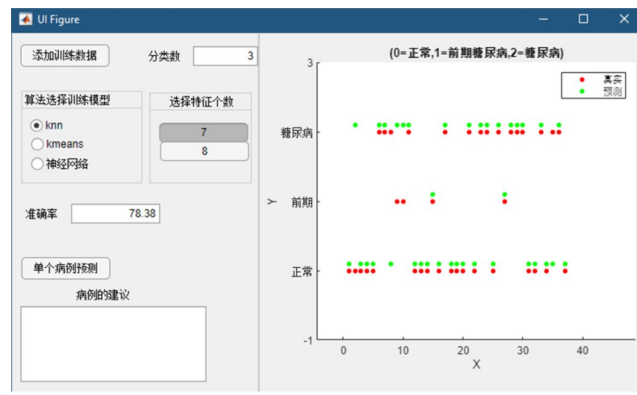


图4 KNN算法的分类准确率(7个输入参数)
Fig.4 Classification accuracy of KNN algorithm (7 input parameters)

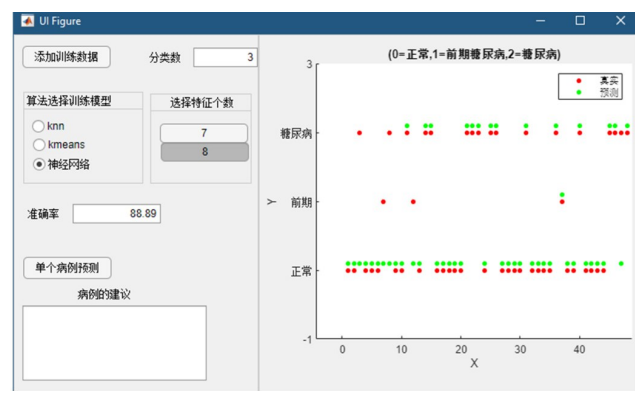


图5 神经网络的分类准确率(8个输入参数)
Fig.5 Classification accuracy of neural network (8 input parameters)

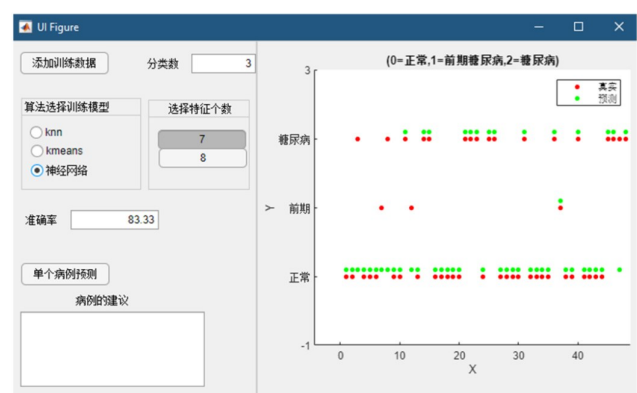


图6 神经网络的分类准确率(7个输入参数)
Fig.6 Classification accuracy of neural network (7 input parameters)

【参考文献】

[1] 关廷均,徐俏娴,尹燕,等. 2型糖尿病患者糖化血红蛋白检测与糖耐量试验的诊断比较[J]. 心电图杂志,2018,7(1): 73.
GUAN T J, XU Q X, YIN Y, et al. Diagnosis comparison of glycosylated hemoglobin detection and glucose tolerance test in patients with type 2 diabetes [J]. Journal of Electrocardiography, 2018, 7 (1): 73.

[2] International Diabetes Federation (IDF). IDF diabetes atlas-7th edition at 2015 world diabetes congress [EB/OL]. [2016-06-29]. <http://www.diabetesatlas.org>.

[3] 王卫庆. 从最新流行病学数据谈糖尿病前期干预的重要性[J]. 药品评价, 2014, 11(13): 18-21.
WANG W Q. Talking about the importance of pre-diabetes intervention from the latest epidemiological data [J]. Drug Evaluation, 2014, 11(13): 18-21.

[4] GANG L. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction [J]. Health Inf Sci Sys, 2016, 4(1): 1-9.

[5] 张俊峰. OGTT和糖化血红蛋白检测在妊娠期糖尿病筛查的应用价值对比分析[J]. 齐齐哈尔医学院学报, 2017, 38(10): 1140-1142.
ZHANG J F. Comparison of different OGTT and HbA1c examination in gestational diabetes screening [J]. Journal of Qiqihar Medical College, 2017, 38(10): 1140-1142.

[6] 屈兰芹,江颖茹. 糖耐量受损对孕产妇及胎儿发育的影响[J]. 中

国医药指南, 2014, 12(35): 20.

QU L Q, JIANG Y R. The influence on maternal and fetal development of pregnant women with glucose tolerance [J]. Guide of China Medicine, 2014, 12(35): 20.

[7] 叶琪. 糖尿病健康管理模式的研究进展综述[J]. 影像研究与医学应用, 2018, 2(3): 1-3.
YE Q. Review on the research progress of diabetes health management mode [J]. Imaging Research and Medical Application, 2018, 2(3): 1-3.

[8] 殷晓. 中医治未病结合健康管理干预治疗糖尿病前期临床研究[J]. 世界最新医学信息文摘, 2019, 19(67): 257.
YIN X. Clinical research on the treatment of prediabetes by TCM combined with health management intervention [J]. World Latest Medical Information, 2019, 19(67): 257.

[9] 梁玉议,张凤清,莫喜萍,等. 气排球运动干预模式在糖尿病前期人群中的应用研究[J]. 右江医学, 2019, 47(9): 673-675.
LIANG Y Y, ZHANG F Q, MO X P, et al. Study on the application of the intervention model of air volleyball in the pre diabetes population [J]. Youjiang Medical Journal, 2019, 47(9): 673-675.

[10] 潘冰莹,陈雄飞,吴雪霁,等. 14种非损伤糖尿病筛查风险评估模型的流行病学特征及其验证[J]. 中国慢性病预防与控制, 2015, 23(1): 1-6.
PAN B Y, CHEN X F, WU X J, et al. Epidemiological characteristic and validation of 14 non-invasive risk score models for screening patients with diabetes [J]. Chinese Journal of Prevention and Control of Chronic Disease, 2015, 23(1): 1-6.

[11] ZHONG X, JING Y, LIN H, et al. A simple tool detected diabetes

- and prediabetes in rural Chinese[J]. J Clin Epidemiol, 2009, 63(9): 1030-1035.
- [12] 韩敏. 人工神经网络基础[M]. 大连: 大连理工大学出版社, 2014: 29-30.
- HAN M. Basis of artificial neural network [M]. Dalian: Dalian University of Technology Press, 2014: 29-30.
- [13] 刘彦杰. 优化BP神经网络在糖尿病患风险分析中的应用神经网络在糖尿病患风险分析中的应用[D]. 兰州: 兰州理工大学, 2016: 11-12.
- LIU Y J. Application of optimized BP neural network in the analysis of the risk of diabetes[D]. Lanzhou: Lanzhou University of Technology, 2016: 11-12.
- [14] 窦小凡. KNN算法综述[J]. 通讯世界, 2018(10): 273-274.
- DOU X F. Overview of KNN algorithm[J]. Telecom World, 2018 (10): 273-274.
- [15] LINDSTROM J, TUOMILEHTO J. The diabetes risk score: a practical tool to predict type 2 diabetes risk [J]. Diabetes Care, 2003, 26(3): 725-731.
- [16] 师正坤, 郭佳, PARRY M, 等. 中国糖尿病风险评估工具的研究现状与进展[J]. 中国全科医学, 2015, 18(20): 2368-2372.
- SHI Z K, GUO J, PARRY M, et al. Research status and progress of Chinese risk assessment tools for diabetes mellitus [J]. Chinese General Practice, 2015, 18(20): 2368-2372.
- [17] 刘媛. 中国糖尿病风险评分表在南昌社区中老年人群糖尿病筛查中的运用及改进[D]. 南昌: 南昌大学, 2018.
- LIU Y. Application and improvement of Chinese diabetes risk score in middle-aged and elderly population of Nanchang community [D]. Nanchang: Nanchang University, 2018.
- [18] 周玉英. 糖化血红蛋白与血脂检测在2型糖尿病检测中临床价值研究[J]. 全科口腔医学电子杂志, 2019, 6(33): 158-159.
- ZHOU Y Y. Study on the clinical value of glycosylated hemoglobin and blood lipid in the detection of type 2 diabetes [J]. Electronic Journal of General Stomatology, 2019, 6(33): 158-159.
- [19] 陈真诚, 杜莹, 邹春林, 等. 基于K-Nearest Neighbor和神经网络的糖尿病分类研究[J]. 中国医学物理学杂志, 2018, 35(10): 1220-1224.
- CHEN Z C, DU Y, ZOU C L, et al. Classification of diabetes based on K-nearest neighbor and neural network [J]. Chinese Journal of Medical Physics, 2018, 35(10): 1220-1224.

(编辑:谭斯允)