

基于变长时间间隔 LSTM 方法的胎儿异常体重预测^①



张硕彦¹, 吴英飞¹, 袁贞明¹, 卢莎², 胡文胜²

¹(杭州师范大学 杭州国际服务工程学院, 杭州 311121)

²(杭州市妇产科医院, 杭州 310008)

通讯作者: 吴英飞, E-mail: wyf@hznu.edu.cn

摘要: 孕妇的产前体检是围产医学的重要组成部分, 产前预测胎儿体重可以为判断胎儿健康发育提供准确的参考. 孕妇的多次体检记录在孕周时间上有不均匀时间间隔分布的特点. 本研究对不均匀时间间隔的处理提出了 LSTM 模型的变种——变长时间间隔的 LSTM 模型 (Variable Time Interval LSTM, VTI-LSTM). 本研究数据来源于 2015~2018 年多家妇产科医院的 10 473 个孕妇的 122 462 条体检记录. 实验比较了传统的公式估算法以及 GBDT, MLP, SVR, RNN, LSTM, VTI-LSTM 等机器学习方法的胎儿体重预测结果, 其中, VTI-LSTM 在低体重和巨大儿的预测上取得良好的预测结果.

关键词: 胎儿体重预测; LSTM; 不均匀时间间隔分布; 公式估算; 机器学习

引用格式: 张硕彦, 吴英飞, 袁贞明, 卢莎, 胡文胜. 基于变长时间间隔 LSTM 方法的胎儿异常体重预测. 计算机系统应用, 2020, 29(3): 39-46. <http://www.c-s-a.org.cn/1003-3254/7334.html>

Fetal Abnormal Weight Prediction Based on Variable Time Interval LSTM

ZHANG Shuo-Yan¹, WU Ying-Fei¹, YUAN Zhen-Ming¹, LU Sha², HU Wen-Sheng²

¹(Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou 311121, China)

²(Hangzhou Women's Hospital, Hangzhou 310008, China)

Abstract: Prenatal physical examination of pregnant women is a important part of perinatal medicine. Prenatal prediction of fetal weight can provide an accurate reference for judging the healthy development of the fetus. The multiple physical examination records have the characteristics of variable time interval distribution during gestational period. This study proposes a variant of LSTM model, Variable Time Interval LSTM (VTI-LSTM), to solve the variable time intervals problem. The data of this study were from 122 462 medical records of 10 473 pregnant women from several women's hospitals during 2015 to 2018. The experiments of fetal weight prediction compare the traditional formula estimation methods with the machine learning methods such as GBDT, MLP, SVR, RNN, LSTM, and VTI-LSTM. The results show that Variable Time Interval LSTM has a good prediction result in the prediction of low birth-weight fetal and macrosomia.

Key words: fetal weight prediction; LSTM; variable time interval distribution; formula estimation; machine learning

1 介绍

预测胎儿体重是产前监护的重要内容, 是医生对孕妇进行临床处理的重要依据. 近年来研究显示, 低体

重儿的存活率和抗感染能力相对低下^[1], 并且与低智商有密切联系^[2]. 而巨大儿则会引起胎儿宫内窘迫、新生儿窒息、肩难产^[3,4]等. 产前预测胎儿体重, 对于避免新

① 基金项目: 国家卫生健康委员会科研项目 (WKJ-ZJ-1911); 杭州市卫生科技计划一般项目 (OO2019054)

Foundation item: Scientific Research Program of National Health Commission of the People's Republic of China (WKJ-ZJ-1911); General Program of Health Science and Technology Plan of Hangzhou Municipality (OO2019054)

收稿时间: 2019-08-08; 修改时间: 2019-09-05, 2019-09-20; 采用时间: 2019-09-23; csa 在线出版时间: 2020-02-28

生儿体重异常,恰当选择分娩方式具有重要意义.孕妇的产前体检记录属于特殊的电子病历.围产医学规定的产前体检、孕妇个人的健康状况和高危妊娠的随访等使得孕妇的体检次数和体检时间间隔各不相同.因而造成了体检事件在妊娠时间上的不均匀时间间隔分布.图1是孕妇产前体检示意图.矩形表示妊娠起点,3个圆点表示孕妇的前3次产前体检记录,倒三角表示分娩.以上事件以时间顺序在数轴上从左往右排列,事件之间的距离大小表示时间间隔的长短.妊娠时间用 ΔT_k 表示,相邻两次体检之间的时间间隔用 Δt_k 表示($k=1,2,3,\dots,N$).孕妇每次体检的时间间隔不仅与妊娠时间有关,而且与相邻两次体检之间的时间间隔有密切联系.传统的RNN和LSTM模型默认序列之间是相等时间间隔的,以序列的先后顺序表示时间信息,在模型层面并没有时间概念,难以充分有效地表征数据不均匀时间间隔的关系.基于以上,本研究从模型层面出发,将妊娠时间 ΔT 和间隔时间 Δt 做嵌入表示,在LSTM模型中的“遗忘门”和“记忆门”分别增添“时间门”,模型以“时间门控”的方式控制“状态”信息,间接表征不均匀时间间隔信息.文章安排如下:第1部分介绍,阐明胎儿体重预测的意义和孕妇的体检事件分布特点,第2部分相关工作,阐述目前胎儿体重预测任务的主流方法和不足,以及深度学习在电子病历中的应用.第3部分是任务定义和模型介绍,第4部分是数据预处理和实验设计,第5部分是实验结果和讨论,最后部分是结论.

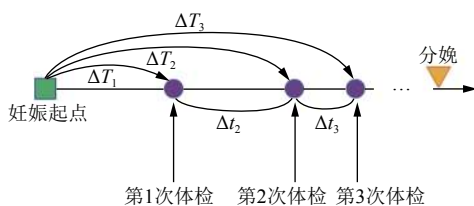


图1 孕妇产前体检记录分布示意图

2 相关工作

在临床工作中,产科医生经常采用宫高腹围测量法^[5-8]和超声测量法^[9,10]估算胎儿体重.两者根据孕妇的检查数据为参数,通过公式来估计胎儿体重.宫高腹围法简单快捷无副作用,孕妇可在家中自行测量.超声法则以其更精确的优点成为目前应用最广泛的方法.但是,这两种方法均存在不同的弊端.首先,由于产科

工作的特殊性,在一些紧急情况下不能及时为孕妇完成超声检查,此时使用宫高腹围法便成为一个很好的选择.超声法是目前较为准确的测量方法,但是,多次做超声检查是否有副作用存在争议,一般情况下,孕妇产前的超声检查仅为3-4次.除以上两种公式估算方法外,机器学习方法也可以将孕妇的宫高、腹围,胎儿的双顶径、头围、腹围、股骨长等数据为参数预测胎儿体重,在刁晓娣等的研究中^[11],使用人工神经网络预测胎儿体重,Xu ZP等^[12]在人工神经网络中添加了正则化条件来预测胎儿体重.虽然,机器学习方法已应用到胎儿体重预测任务中,但是其方法大多为人工神经网络,不同的研究虽对人工神经网络有不同的修改,但其本质仍旧是MLP(多层感知机).无论是宫高腹围法、超声测量法还是经典机器学习方法,仅取孕妇分娩前一周内的数据作为估算或预测参数,而忽略了孕妇在多次体检记录中反映出来的变化.其次,在胎儿体重预测的研究中,尚未将孕妇的高危妊娠和疾病信息考虑进去.因此,此类研究不具有普适性.

3 变长时间间隔 LSTM 模型

孕妇在孕周时间上的体检统计分布如图2所示,横轴表示妊娠周,深色的小矩形表示孕妇在对应的妊娠周做了体检,每一横行表示孕妇体检周的一种统计分布,纵轴表示每一种统计分布的人数.例如,孕妇a的体检时间周是序列[16, 20, 24, 28, 30, 32, 34, 35, 36, 37, 38, 39],序列中的每一项是做体检的妊娠周,这串序列为一种体检周统计分布,假设孕妇b与孕妇a的体检时间周完全一致,即,孕妇b的体检妊娠周也为[16, 20, 24, 28, 30, 32, 34, 35, 36, 37, 38, 39],那么,符合该统计分布的孕妇的数量加1,反映在图2中是最上方的第一横行,横坐标为[16, 20, 24, 28, 30, 32, 34, 35, 36, 37, 38, 39]处是深色矩形.每一横行在左边对应的纵坐标上的数值是体检的妊娠周完全符合该统计分布的孕妇人数,即,体检孕周是[16, 20, 24, 28, 30, 32, 34, 35, 36, 37, 38, 39]统计分布的孕妇共有21人.10 473个孕妇共有9480种统计分布.按照每种分布的孕妇人数降序排列,图2绘制了人数最多的前30种统计分布.根据图2可知,在9480种体检统计分布中,人数最多的仅有21人,由于大部分孕妇的体检周的分布差异巨大,因此,孕妇的体检记录在妊娠周时间上无法对齐,需要在传统的LSTM中引入时间信息.

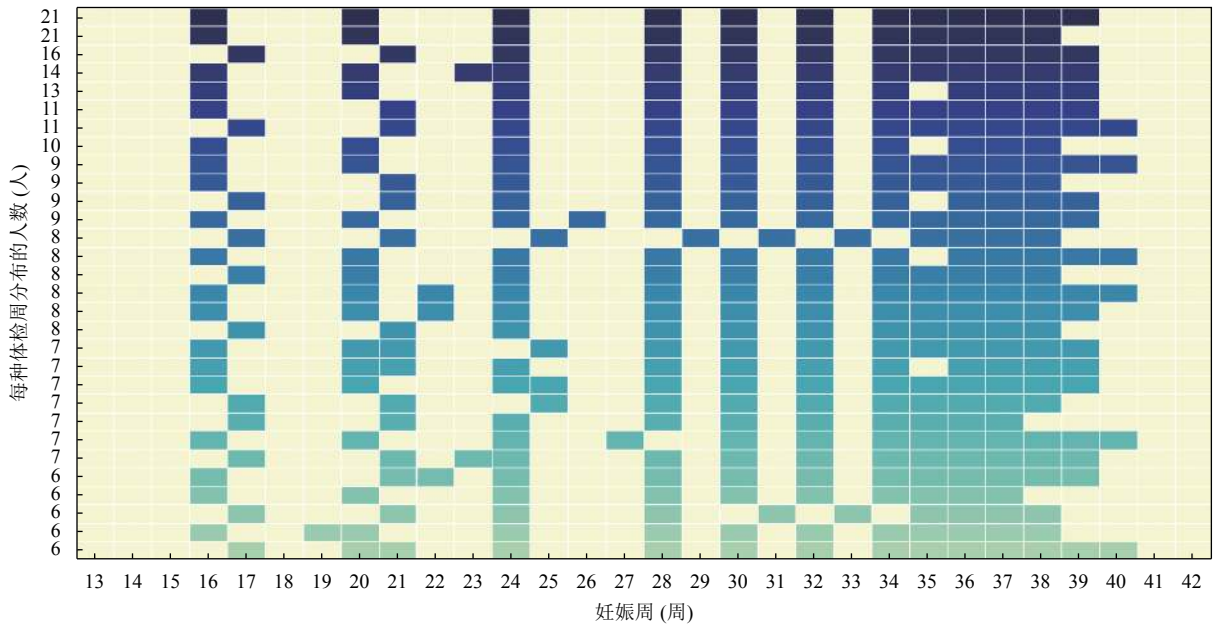


图2 孕妇体检周统计分布图(前30种统计分布)

本文用 $P = \{p^1, p^2, \dots, p^i, \dots, p^N\}$ 表示孕妇集合, 用 $R = \{r^1, r^2, \dots, r^i, \dots, r^N\}$ 表示孕妇体检记录集合. N 表示孕妇总人数, i 表示第 i 个孕妇, 孕妇 p^i 对应的体检记录为 r^i . 每一个孕妇的体检记录和体检时间表示为 $M^i := [(r_1^i, t_1^i), (r_2^i, t_2^i), \dots, (r_k^i, t_k^i), \dots, (r_l^i, t_l^i)]$. 其中, r_k^i 和 t_k^i 分别表示第 i 个孕妇的第 k 次体检的记录和时间, l 表示体检次数. 根据 $[t_1^i, t_2^i, \dots, t_k^i, \dots, t_l^i]$, 当前的妊娠时间 ΔT_k^i 和距上次体检的间隔时间 Δt_k^i 的计算方式分别如式 (1), 式 (2):

$$\Delta T_k^i = t_k^i \quad (1)$$

$$\Delta t_k^i = t_k^i - t_{k-1}^i, \text{ s.t. } k \in N^+, t_0^i = 0 \quad (2)$$

本文的目标是根据 M^i , 输出胎儿出生体重的预测值. 在标准的 RNN 和 LSTM 中, 第 k 个时间步的输入是 $x_k = r_k^i \oplus t_k^i$, 其中, \oplus 表示 r_k^i 和 t_k^i 做向量拼接. 在变长时间间隔 LSTM 中, 第 k 个时间步的输入是 $x_k, \Delta T_k^i, \Delta t_k^i$. 由于孕妇第一次体检无间隔时间 Δt , 故令 $\Delta t_1^i = \Delta T_1^i$.

变长时间间隔 LSTM 模型即为 Variable Time Interval-LSTM (VTI-LSTM). 标准的 LSTM 模型的表达式^[13]如式 (3)–式 (7):

$$i_k = \sigma_i(x_k W_{xi} + h_{k-1} W_{hi} + w_{ci} \odot c_{k-1} + b_i) \quad (3)$$

$$f_k = \sigma_f(x_k W_{xf} + h_{k-1} W_{hf} + w_{cf} \odot c_{k-1} + b_f) \quad (4)$$

$$c_k = f_k \odot c_{k-1} + i_k \odot \tanh_c(x_k W_{xc} + h_{k-1} W_{hc} + b_c) \quad (5)$$

$$o_k = \sigma_o(x_k W_{xo} + h_{k-1} W_{ho} + w_{co} \odot c_k + b_o) \quad (6)$$

$$h_k = o_k \odot \sigma_h(c_k) \quad (7)$$

LSTM 模型结构如图 3(a) 所示. 我们将妊娠时间 ΔT 和间隔时间 Δt 做嵌入表征^[14], 如式 (8), 式 (9):

$$T1_k = \sigma_{t1}(x_k W_{xT} + \sigma_{tt1}(\Delta T_k W_{tt1}) + b_{t1}) \quad (8)$$

$$T2_k = \sigma_{t2}(x_k W_{xt} + \sigma_{tt2}(\Delta t_k W_{tt2}) + b_{t2}) \quad (9)$$

然后, 本文分别向 Standard LSTM 模型的“输入门”和“遗忘门”中以时间门控的方式引入妊娠时间 ΔT 和间隔时间 Δt , 即, 将式 (5) 换为式 (10):

$$c_k = (1 - T2_k) \odot f_k \odot c_{k-1} + i_k \odot T1_k \odot \tanh_c(x_k W_{xc} + h_{k-1} W_{hc} + b_c) \quad (10)$$

对于式 (10) 中 $(1 - T2_k)$ 项, 时间间隔的嵌入表征 $T2_k$ 越小, 模型对过去事件的“遗忘”越少. 而 $T1_k$ 则表示某次体检的妊娠时间对该次输入的“记忆”的多少. VTI-LSTM 模型结构如图 3(b) 所示, 左上角分别是两个“时间门”时间门 1 和时间门 2, 时间门 1 的输入是 x_k 和 ΔT_k , 时间门 2 的输入是 x_k 和 Δt_k .

4 实验和数据

4.1 特征筛选

研究表明, 胎儿体重不仅与孕妇的宫高^[15]、腹围^[16]相关, 而且与身高体重^[17,18]、胎次^[19]、年龄、糖尿病^[20]

等因素有关. 本研究中, 本文选取孕妇的宫高腹围和高危因素等方便获取的参数作为模型的输入特征. 根据

数据中的特征, 计算各项特征与胎儿体重的皮尔森相关系数和 P 值. 皮尔森相关系数计算公式如式 (11).

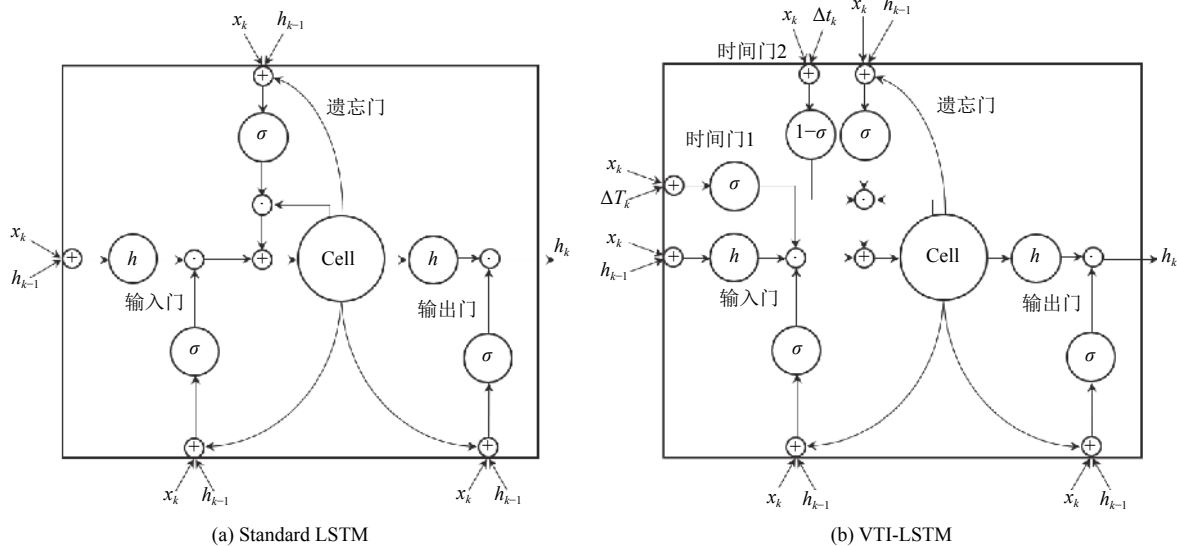


图3 两个 LSTM 模型图

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \quad (11)$$

其中, x_i 和 y_i 分别是两个变量的数据项, \bar{x} 和 \bar{y} 分别是两个变量的平均值, N 是样本总数. s_x 和 s_y 分别是两个变量的标准差. P 值即为原假设 $H_0: \rho = 0$, 备择假设 $H_1: \rho \neq 0$ 的 t 分布双边检验. P 值计算公式如式 (12), 式 (13):

$$t = \frac{\rho \sqrt{n-2}}{\sqrt{1-\rho^2}} \quad (12)$$

$$P_{\text{value}} = 2 \times P(T > t) \quad (13)$$

其中, T 服从 $n-2$ 自由度的 t 分布, P_{value} 即为 P 值. 计算结果如表 1 所示, 保留 3 位有效数字. 我们选择 P 值低于 0.05 的特征, 包含妊娠周、年龄、身高、体重、宫高、腹围、胎次、早产史、胎盘异常、胎位不正、胎膜早破、糖尿病、辅助生殖、多胎共 14 个特征. 其中, 妊娠周、年龄、身高、体重、宫高、腹围、胎次属于数值型数据, 早产史、胎盘异常、胎位不正、胎膜早破、糖尿病、辅助生殖、多胎属于类别型数据.

表 1 相关性分析结果

数据项	妊娠周	年龄	身高	体重	宫高	腹围	胎心率	血型	RH 型	婚姻状况	胎次	早产史
系数	0.526	0.019	0.159	0.276	0.431	0.345	-0.019	-0.008	-0.005	0.001	0.051	-0.020
P 值	0.000	0.046	0.000	0.000	0.000	0.000	0.058	0.415	0.629	0.951	0.000	0.040
数据项	疤痕子宫	甲状腺病	胎盘异常	胎位不正	胎膜早破	高血压	糖尿病	心脏病	性病	辅助生殖	多胎	
系数	-0.001	-0.005	-0.020	-0.027	-0.046	0.011	0.037	0.016	0.004	-0.072	-0.214	
P 值	0.936	0.601	0.041	0.005	0.000	0.243	0.000	0.100	0.715	0.000	0.000	

4.2 数据预处理

将胎儿体重以小于 2500 g、在 2500 g 和 4000 g 之间、大于 4000 g 为标准分别划分为低体重儿、正常体重儿、巨大儿, 如表 2 所示, 可以看出, 低体重儿和巨大儿样本量与正常体重儿数量差距较大, 因此在分类任务的训练集中, 将低体重儿和巨大儿样本做过采

样处理, 在总体的回归任务上不做过采样处理. 为避免数据中不同的量纲对模型造成影响, 本文采用归一化的方式处理数值型数据. 绘制胎儿体重分布图, 如图 4 所示, 横坐标表示胎儿体重, 纵坐标表示与该体重对应的胎儿数量. 根据横坐标间距和样本数量, 将样本划分为 3 段, 分别用倒三角、圆点、正三角表示. 观察可得,

胎儿体重大多集中在中间圆点的区域,两边倒三角和正三角区域则较为稀疏,并且胎儿体重在横坐标上跨度较大,使用常规的归一化方法必然导致中间区域大量样本的归一化结果差异非常小,导致模型表现较差.因此,我们采用分段缩放的方法,牺牲小样本上横坐标的间距,放大大量样本的横坐标间距.倒三角样本点、圆样本点、正三角样本点分布的横坐标区间分别记为 I_1 、 I_2 、 I_3 , 区间内胎儿体重数值种类的数量分别记为 C_1 、 C_2 、 C_3 , 则各区间的密度分别记为式 (14)~式 (16),

表2 胎儿体重分类

胎儿体重 weight (g)	数量 (个)
weight<2500	357
2500<=weight<=4000	9774
weight>4000	342

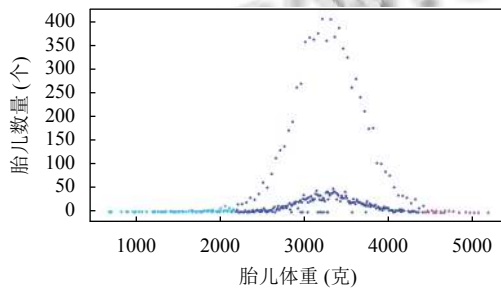


图4 胎儿体重分布图

$$D_1 = \frac{C_1}{I_1} \quad (14)$$

$$D_2 = \frac{C_2}{I_2} \quad (15)$$

$$D_3 = \frac{C_3}{I_3} \quad (16)$$

在区间 I_1 上,对样本进行一次线性归一化,如式 (17):

$$V_{new}^{I_1} = \frac{V_{origin}^{I_1} - V_{min}^{I_1}}{V_{max}^{I_1} - V_{min}^{I_1}} \quad (17)$$

其中, $V_{origin}^{I_1}$ 是区间 I_1 上原始数据的取值, $V_{min}^{I_1}$ 和 $V_{max}^{I_1}$ 分别是区间中最小值和最大值. $V_{new}^{I_1}$ 是经过公式计算后的结果.对 I_2 和 I_3 也做类似计算得 $V_{new}^{I_2}$ 和 $V_{new}^{I_3}$. 那么,各区间分段缩放后的取值分别为式 (18)~式 (20),

$$I_{1new} = D_1 \times V_{new}^{I_1} \quad (18)$$

$$I_{2new} = D_2 \times V_{new}^{I_2} + D_1 \quad (19)$$

$$I_{3new} = D_3 \times V_{new}^{I_3} + D_1 + D_2 \quad (20)$$

将 I_{1new} 、 I_{2new} 、 I_{3new} 合并为一个大区间,记为 I_{total} ,再对 I_{total} 做一次类似式 (14) 的线性归一化计算.

经过上述分段缩放和归一化计算,最终,数值类型数据的间隔相对均匀,且保留了各子区间内的线性关系,可提升模型效果.我们对身高、体重、宫高、腹围、胎儿体重 5 个数值型数据做了上述计算,其他数值型数据分布较为均匀,仅做简单的线性归一化.对于类别型数据,本文采用 one-hot 的编码方式,当某一孕妇的体检记录包含多种类别型数据,那么与该孕妇对应的类别信息为 multi-hot 向量.

孕妇的体检次数如图 5 所示,横坐标表示孕妇体检次数,纵坐标表示孕妇人数.最少体检次数为 1 次,最多体检次数为 24 次,孕妇的体检记录是一个变长序列.对于时序模型,需要将孕妇的体检记录做填充.上面提到的特征筛选中,孕妇的一次体检记录共有 14 维.本文用 1×14 的 0 向量将孕妇的体检记录填充至 24 次,0 向量不参与模型计算.因此,10 473 个孕妇的 122 462 条体检记录转化为一个三维矩阵,第一维度表示最大体检次数,第二维度表示每条体检记录的特征数量,第三维度表示孕妇总人数.

4.3 实验设置

本文将产前的体检记录作为模型的输入,胎儿出生体重作为预测目标,10 473 个孕妇的最后一次体检记录与分娩时间在同一周内.实验分为公式法和机器学习方法 2 部分.公式法是选取 4 个不同的胎儿体重估算公式^[5-8]作为对照,4 个估算公式如式 (21)~式 (24),分别对应公式法 1~公式法 4.其中, BW 是胎儿的出生体重(单位为 g), FH 是宫高(单位为 cm), AG 是腹围(单位为 cm).

$$BW = -1.515 + 0.092 \times FH + 0.016 \times AG \quad (21)$$

$$BW = 13.3 \times FH - 943 \quad (22)$$

$$BW = (AG - 67.87) / 7.2 \quad (23)$$

$$BW = -2700 + 123 \times FH + 20 \times AG \quad (24)$$

机器学习方法采用 GBDT、MLP、SVR、RNN、LSTM、VTI-LSTM 等 6 种模型,损失函数采用二次代价函数,如式 (25):

$$L = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (25)$$

其中, y_i 和 \hat{y}_i 分别是预测值和真实值. N 是训练集中孕妇总人数.为直观展示预测误差,我们采用平均相对误差

如式 (26):

$$MRE = \frac{1}{M} \sum_{j=1}^M \frac{|y_j - \hat{y}_j|}{\hat{y}_j} \times 100\% \quad (26)$$

其中, y_j 是预测值, \hat{y}_j 是真实值. M 是测试集中孕妇总人数. 在实际临床工作中, 我们更关注异常体重胎儿, 仅以回归任务的相对误差作为预测结果显然是不够的, 因此, 实验分为两个任务:

(1) 对低体重儿、正常体重儿、巨大儿的分类. 分

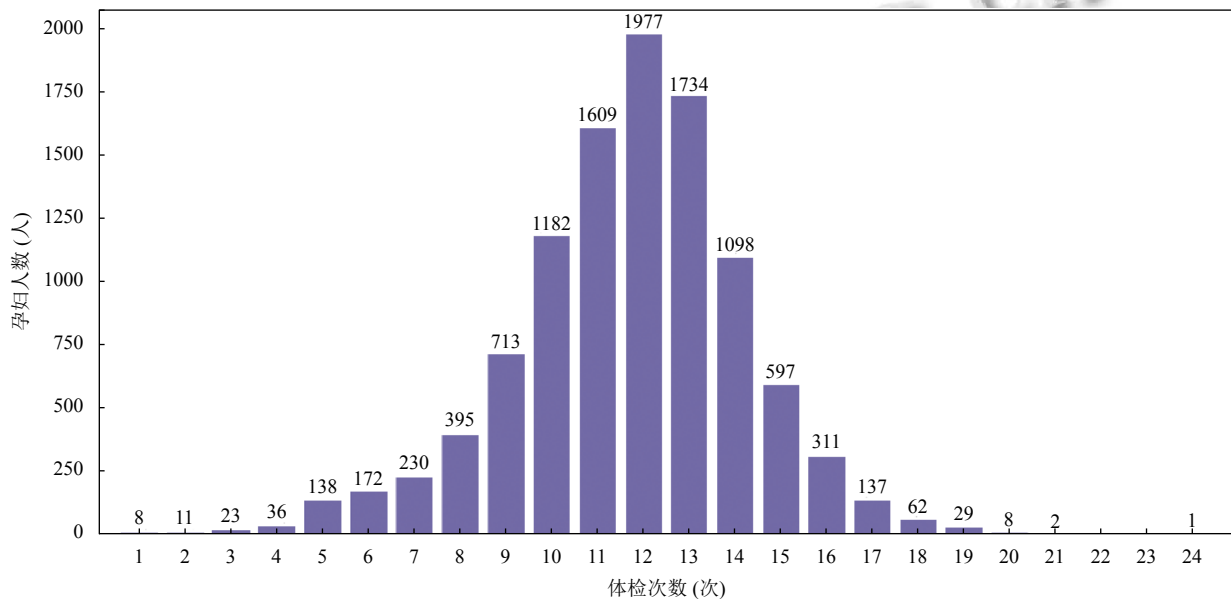


图5 孕妇体检次数分布图

5 实验结果和讨论

胎儿体重预测结果的平均相对误差如表3所示. 其中, 实验结果的前3列是分类任务在低体重儿、正常体重儿、巨大儿的预测上的MRE. 第4列是总体回归任务的误差, 也即, 前3列的机器学习方法中, 对训练集中的小样本做了过采样处理, 第4列的机器学习方法中, 对训练集中的小样本没有做过采样处理, 保留原始数据中各类样本的比例. 公式法无训练集, 直接在测试集上计算即可. 总体来看4个公式法预测的结果弱于机器学习方法. 式(21)、式(24)方法包含宫高腹围两个参数, 式(22)仅有宫高参数, 式(23)仅有一个腹围参数, 所以, 式(23)法容易受到孕妇腹部脂肪的干扰, 结果较差, 式(22)相对式(23)较好, 但弱于式(21)和式(24). 式(21)的预测结果偏向于正常体重儿, 在所有的方中取得最小的正常体重儿预测结果误差. GBDT、MLP、SVR等3种经典的机器学习模型的预

类按照表2划分为3类, 我们从低体重儿、正常体重儿、巨大儿中各抽取50个作为分类任务的测试集, 根据分类预测结果计算各类体重的MRE并绘制混淆矩阵. 由于胎儿体重的各类样本数量不平衡, 因此需要对训练集中的低体重儿和巨大儿做过采样处理.

(2) 计算总体上的回归误差, 训练集不做过采样处理, 总体回归的测试集保持数据中原始的各类体重儿的比例. 训练集和测试集按照10:1划分.

测结果并无较大差异. RNN、LSTM、VTI-LSTM等3种时序模型方法中, LSTM结果仅次于VTI-LSTM, RNN结果较差, 说明仅使用简单的递归循环难以学习到孕妇多次体检的变化. VTI-LSTM在低体重儿和巨大儿MRE上取得最好的预测结果, 而在正常体重儿MRE上弱于式(21).

由于在分类任务中, 我们对小样本做了过采样处理, 所以导致机器学习模型牺牲正常体重儿的结果, 偏向低体重儿和巨大儿. 观察第4列的结果, 不做小样本过采样处理的总体回归误差仍旧是VTI-LSTM效果最好, 根据表2的统计结果, 占据样本大多数的正常体重儿的体重范围是2500~4000g, 区间跨度并不大, 所以导致机器学习模型的总体MRE结果相差较小. 由于我们采用的是MRE作为误差度量的方式, 根据图4可以看出, 低体重儿的多样性比巨大儿丰富, 并且低体重儿

计算 MRE 公式的分母较小,所以在所有的方法中,低体重儿的 MRE 高于巨大儿。

表3 测试集中胎儿体重预测的平均相对误差 MRE(%)

模型和方法	低体重儿	正常体重儿	巨大儿	总体
公式法 1-式 (21) ^[5]	44.72	6.62	16.77	8.37
公式法 2-式 (22) ^[6]	62.79	15.97	9.80	17.46
公式法 3-式 (23) ^[7]	94.08	30.81	22.76	34.28
公式法 4-式 (24) ^[8]	55.29	9.54	8.76	11.81
GBDT	17.19	9.77	9.31	7.84
MLP	21.77	9.62	9.16	8.96
SVR	19.95	10.41	9.95	7.83
RNN	55.56	13.04	15.56	8.56
LSTM	14.17	9.32	7.65	8.18
VTI-LSTM	13.87	10.88	7.00	7.80

分类预测结果的混淆矩阵如图6所示。纵坐标表示真实类别,横坐标表示预测类别。L表示低体重儿(low birth-weight infant),N表示正常体重儿(normal birth-weight infant),M表示巨大儿(macrosomia)。公式法1、公式法2、公式法4中,将大部分样本分类到正常体重中,仅使用腹围作为预测参数的公式法3则倾向于将样本分类到巨大儿中,预测结果受到了孕妇腹部脂肪的影响。GBDT、MLP、SVR三者中,MLP结果相对较差,GBDT和SVR无明显差异。RNN的分类结

果与公式法接近,LSTM和VTI-LSTM明显优于其他方法,LSTM与VTI-LSTM相比,LSTM偏向于分类到正常体重儿中,VTI-LSTM则在低体重儿和巨大儿的分类中取得最好的结果,说明不均匀离散时间在“输入门”和“遗忘门”上的嵌入表征起到了很大的作用。上述的方法和模型中,无论是哪一种,对低体重儿和巨大儿的分类仍旧是一个挑战。根据表2中的数据,低体重儿和巨大儿分别仅占新生儿总数的3.41%和3.27%,各类样本比例失衡,虽然在训练集中我们采取了对小样本过采样的预处理方式,但仍难以避免小样本数据缺乏多样性的问题。另外,体重低于2000g的低体重儿和体重高于4500g的巨大儿分别占低体重儿和巨大儿较小的比例,导致低体重儿、正常体重儿、巨大儿的分类间隔较小。对比低体重儿和巨大儿的分类,低体重儿的分类结果明显优于巨大儿。原因有如下两点:第一,在我们的特征中,包含孕周信息,分娩低体重儿的孕妇的妊娠时间普遍少于正常的妊娠时长。第二,由于现代社会生活水平普遍较高,胎儿体重偏大,在孕36周时,许多胎儿已经发育成熟,且体重在正常范围内,这类胎儿和足月儿相比会提前1-2周分娩,增加了模型识别巨大儿的难度。

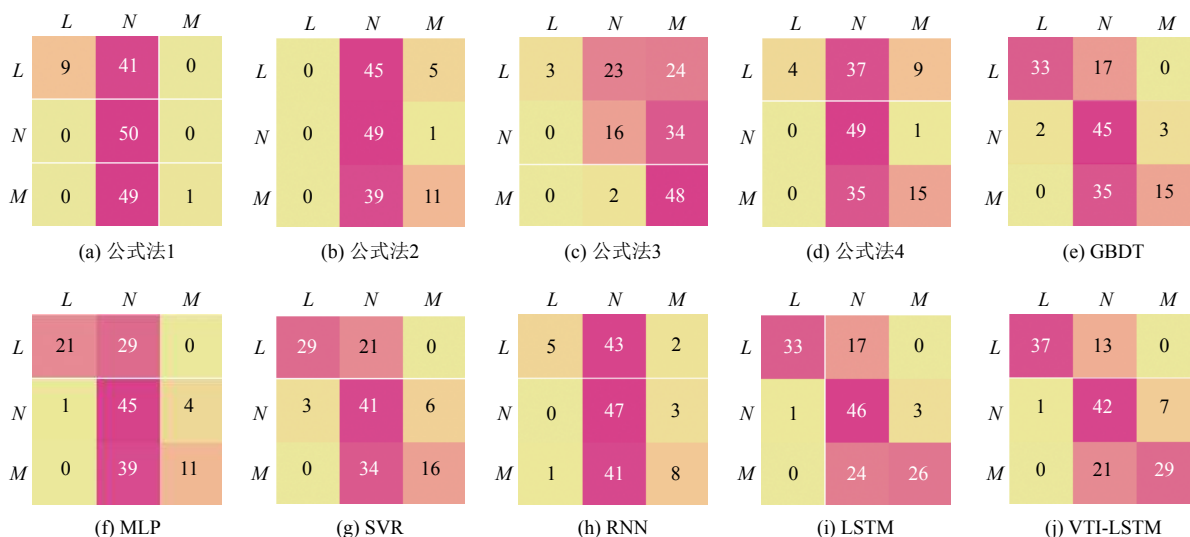


图6 测试集上预测结果的混淆矩阵

6 结论

本研究在胎儿体重预测任务上分别使用了4种公式法和3种经典的机器学习方法以及3种时序模型方法。其中,本研究提出的VTI-LSTM模型在低体重儿和

巨大儿的分类预测中取得了较好的结果并在总体的误差回归上取得最小的MRE。传统的公式法弱于经典的机器学习模型和时序模型。而时序模型中的LSTM可以学习到孕妇每次体检的变化以及胎儿的生长速率,

预测结果有大幅提升. VTI-LSTM 将每次体检的时间间隔和体检的妊娠时间在模型层面上表征, 模型可以学习到孕妇体检记录的不均匀时间间隔, 又在 LSTM 的基础上得到提升. 综上, 本研究可为医生和孕妇判断胎儿生长发育提供一个相对准确的参考.

参考文献

- 1 McIntire DD, Bloom SL, Casey BM, *et al.* Birth weight in relation to morbidity and mortality among newborn infants. *The New England Journal of Medicine*, 1999, 340(16): 1234–1238. [doi: [10.1056/NEJM199904223401603](https://doi.org/10.1056/NEJM199904223401603)]
- 2 Matte TD, Bresnahan M, Begg MD, *et al.* Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: Cohort study. *British Medical Journal*, 2001, 323(7308): 310–314. [doi: [10.1136/bmj.323.7308.310](https://doi.org/10.1136/bmj.323.7308.310)]
- 3 Modanlou HD, Dorchester WL, Thorosian A, *et al.* Macrosomia—Maternal, fetal, and neonatal implications. *Obstetrics and Gynecology*, 1980, 55(4): 420–424.
- 4 Oral E, Çağdaş A, Gezer A, *et al.* Perinatal and maternal outcomes of fetal macrosomia. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 2001, 99(2): 167–171.
- 5 Woo JSK, Ngan HYS, Au KKL, *et al.* Estimation of fetal weight in utero from symphysisfundal height and abdominal girth measurements. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 1985, 25(4): 268–271. [doi: [10.1111/j.1479-828X.1985.tb00742.x](https://doi.org/10.1111/j.1479-828X.1985.tb00742.x)]
- 6 Ghaemmaghami F, Jamal A, Soleimani MR, *et al.* Parturient fundal height and birth weight estimation. *Archives of Iranian Medicine*, 2002, 5(2): 80–83.
- 7 Mohanty C, Das BK, Mishra OP. Parturient abdominal circumference as a predictor of low birthweight. *Journal of Tropical Pediatrics*, 2000, 46(6): 363–364. [doi: [10.1093/tropej/46.6.363](https://doi.org/10.1093/tropej/46.6.363)]
- 8 凌萝达. 难产理论与实践. 重庆: 重庆出版社, 2006.
- 9 Hadlock FP, Harrist RB, Carpenter RJ, *et al.* Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. *Radiology*, 1984, 150(2): 535–540. [doi: [10.1148/radiology.150.2.6691115](https://doi.org/10.1148/radiology.150.2.6691115)]
- 10 Shepard MJ, Richards VA, Berkowitz RL, *et al.* An evaluation of two equations for predicting fetal weight by ultrasound. *American Journal of Obstetrics and Gynecology*, 1982, 142(1): 47–54. [doi: [10.1016/S0002-9378\(16\)32283-9](https://doi.org/10.1016/S0002-9378(16)32283-9)]
- 11 刁晓娣, 江志斌, 刘瑾. 根据孕妇参数预测胎儿体重的神经网络方法. *中国生物医学工程学报*, 1999, 18(2): 155–158. [doi: [10.3969/j.issn.0258-8021.1999.02.006](https://doi.org/10.3969/j.issn.0258-8021.1999.02.006)]
- 12 Xu ZP, Shen AF. A new model of estimating fetal macrosomia based on neural network. *Proceedings of 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science*. Hong Kong, China. 2010. 689–692.
- 13 Graves A. Generating sequences with recurrent neural networks. arXiv: 1308.0850, 2013.
- 14 Zhu Y, Li H, Liao YK, *et al.* What to do next: Modeling user behaviors by time-LSTM. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. Melbourne, Australia. 2017. 3602–3608.
- 15 Bothner BK, Gulmezoglu AM, Hofmeyr GJ. Symphysis fundus height measurements during labour: A prospective, descriptive study. *African Journal of Reproductive Health/LaRevue Africaine de la Santé Reproductive*, 2000, 4(1): 48–55.
- 16 Smith GCS, Smith MFS, McNay MB, *et al.* The relation between fetal abdominal circumference and birthweight: Findings in 3512 pregnancies. *BJOG: An International Journal of Obstetrics and Gynaecology*, 1997, 104(2): 186–190. [doi: [10.1111/j.1471-0528.1997.tb11042.x](https://doi.org/10.1111/j.1471-0528.1997.tb11042.x)]
- 17 Viswanathan M, Siega-Riz AM, Moos MK, *et al.* Outcomes of maternal weight gain. *Evidence Report/technology Assessment*, 2008, (168): 1–223.
- 18 Yu CKH, Teoh TG, Robinson S. Obesity in pregnancy. *BJOG: An International Journal of Obstetrics and Gynaecology*, 2006, 113(10): 1117–1125. [doi: [10.1111/j.1471-0528.2006.00991.x](https://doi.org/10.1111/j.1471-0528.2006.00991.x)]
- 19 Shah PS. Parity and low birth weight and preterm birth: A systematic review and meta-analyses. *Acta obstetrica et gynecologica Scandinavica*, 2010, 89(7): 862–875. [doi: [10.3109/00016349.2010.486827](https://doi.org/10.3109/00016349.2010.486827)]
- 20 HAPO Study Cooperative Research Group. Hyperglycemia and adverse pregnancy outcomes. *New England Journal of Medicine*, 2008, 358(19): 1991–2002. [doi: [10.1056/NEJMoa0707943](https://doi.org/10.1056/NEJMoa0707943)]