

基于 LSTM 与 XGBOOST 混合模型的孕妇产后出血预测^①



周彤彤^{1,2}, 俞凯^{1,2}, 袁贞明^{1,2}, 卢莎³, 胡文胜³

¹(杭州师范大学 信息科学与工程学院, 杭州 311121)

²(移动健康管理教育部工程研究中心, 杭州 311121)

³(杭州市妇产科医院, 杭州 310008)

通讯作者: 俞凯, E-mail: yk@hznu.edu.cn

摘要: 孕妇产后大出血是造成全球孕妇死亡的重要因素之一, 在我国位居孕妇死亡原因首位, 然而对产后出血的提前判定一直以来都是医学上一个难题. 电子病历的普及, 以及机器学习和深度学习技术的发展, 为预知孕妇产后大出血提供了基于大数据的解决办法. 本文提出利用孕妇的电子病历数据, 构建基于 LSTM 和 XGBoost 的混合模型来预测孕妇产后大出血. 实验结果表明, 利用基于 LSTM 和 XGBoost 的混合模型对孕妇产后大出血进行预测是可行的, 能够为医生判断孕妇产后出血情况提供参考, 为孕妇分娩时是否需要备血方案提供决策支持, 对降低产后大出血致死率具有积极意义.

关键词: 产后出血; eXtreme Gradient Boosting (XGBoost); Long Short-Term Memory (LSTM); 机器学习; 深度学习

引用格式: 周彤彤, 俞凯, 袁贞明, 卢莎, 胡文胜. 基于 LSTM 与 XGBOOST 混合模型的孕妇产后出血预测. 计算机系统应用, 2020, 29(3): 148-154. <http://www.c-s-a.org.cn/1003-3254/7316.html>

Predictive Analysis of Postpartum Hemorrhage Based on LSTM and XGBoost Hybrid Model

ZHOU Tong-Tong^{1,2}, YU Kai^{1,2}, YUAN Zhen-Ming^{1,2}, LU Sha³, HU Wen-Sheng³

¹(School of Information Science and Engineering, Hangzhou Normal University, Hangzhou 311121, China)

²(Engineering Research Center of Mobile Health Management System (Ministry of Education), Hangzhou 311121, China)

³(Hangzhou Women's Hospital, Hangzhou 310008, China)

Abstract: Postpartum hemorrhage in pregnant women is one of the most important factors of maternal death around the world, ranking first in China. However, the early diagnosis of postpartum hemorrhage has always been a medical problem. With the popularity of Electronic Health Records and the development of machine learning and deep learning technologies, new solutions have been provided for predicting postpartum hemorrhage in pregnant women. This study proposes to construct a mixed prediction model of postpartum hemorrhage based on LSTM and XGBoost by using the Electronic Health Records of pregnant women. The experimental results show that the hybrid model based on LSTM and XGBoost is feasible to predict postpartum hemorrhage in pregnant women. It can provide a reference for doctors to judge the situation of postpartum hemorrhage and provide decision-making support for whether blood preparation would be needed during delivery. It is of positive significance to reduce the mortality rate of postpartum hemorrhage.

Key words: postpartum hemorrhage; eXtreme Gradient Boosting (XGBoost); Long Short-Term Memory (LSTM); machine learning; deep learning

① 基金项目: 国家卫生健康委员会科研项目-浙江省卫生健康重大项目 (WKJ-ZJ-1911); 卫生科技计划一般项目 (OO2019054)

Foundation item: Project of National Health Commission-Major Project of Public Health of Zhejiang Province (WKJ-ZJ-1911); General Science and Technology Program of Public Health (OO2019054)

收稿时间: 2019-07-28; 修改时间: 2019-09-02; 采用时间: 2019-09-18; csa 在线出版时间: 2020-02-28

引言

联合国千年发展目标中,第五个议题是关于改善孕产妇的健康问题.报告显示,过去25年期间,孕产妇死亡率仍高达约十万分之四十五^[1],孕妇产后大出血是造成全球孕妇死亡的重要因素之一,在我国居首位.研究显示,在经济发达的国家和地区,产妇产后出血的概率会显著提升.随着经济与社会的发展,孕产妇健康问题成为人们关心的热点问题.据调研,某东部医院一年中39.09%的孕妇有产后出血的情况,4.13%的产妇出现产后大出血情况.产后出血一旦发生,预后严重,若持续时间较长、休克较重,即便获救,仍有可能发生严重并发症,因此产后出血的防治工作意义重大^[2].

然而,现有医学手段难以直接对孕妇产后大出血进行精准的预测.目前,医学上无法准确判定产后大出血的成因,临床上医生只能根据某些症状如胎盘前置、妊高症等经验判断产妇产后大出血的可能性.随着医疗信息化的发展,孕产妇的基本生理特征、历次孕检信息都被记录在电子病历中,这为医生全面掌握孕妇产后身体状况,预判产后出血风险提供了便利.然而,由于产后出血相关因素众多且联系复杂,即使是经验丰富的医生,面对这些体量巨大、类型异构、内部关联复杂的临床大数据,都难以全面客观地进行分析^[3],对缺乏经验的年轻医生更是一个巨大的挑战.随着机器学习、深度学习的发展,借助计算机手段挖掘电子病历数据之间的隐藏关系^[4],建立疾病风险预测模型,为复杂的疾病诊断提供支持,成为近年来计算机与医学领域研究的热点.本文尝试利用医院电子病历中孕妇产后基本数据以及历次孕检数据结合机器学习方法,实现对孕妇产后大出血的预测.

目前,医学上常利用传统机器学习或深度神经网络构建疾病预测模型.深度模型对于时间序列数据有良好的处理能力,然而深度模型难以解释,医学上对于模型的可解释性有较高的要求,传统树模型具有良好的可解释性^[5],通过回溯树的结构能够探究出决策形成的原因,为医生诊断提供参考.

孕检的电子病历数据中包含产后大出血相关的特征,孕检数据中孕周的变化是有关联的时间序列数据.传统树模型难以挖掘时间序列中隐含的信息,LSTM(长短期记忆网络)模型^[6]是一种时间递归神经网络,适合处理和预测时间序列中间隔和延迟相对较

长的重要事件.XGBoost模型是一种优秀的集成树模型,具有高效、灵活、鲁棒的特点.为结合两者优点,本文利用LSTM和XGBoost构建产后大出血混合预测模型.

1 数据集处理

模型用到的孕妇数据分为孕妇基础信息和孕检信息两部分,孕妇基础信息包含孕妇自身基本信息,如孕妇的身高、体重、年龄、过往孕史、家庭遗传病史等.孕检信息分为首次孕检和历次孕检,首次孕检信息存储孕妇进行第一次孕检的结果,历次孕检包含孕妇在孕期内每次孕检的结果,主要包含胎儿的生长情况、孕妇的体征指标,如血压、血糖、腹围等.

1.1 数据预处理

数据预处理是提高数据质量,避免数据缺失、错误对模型性能影响的重要步骤.本文中,数据预处理主要分3部分:缺失值处理、归一化处理以及孕检数据的标准化.

孕检数据中存在各孕妇孕检项目不同或未记录检查结果的情况,需进行缺失值处理.根据实际情况,孕检结果中未记录数据为该项目的正常值,另外孕检结果的众数或平均值也为该项目的正常值.因此,在实验中缺失的离散值、连续值分别采用其他孕妇同一孕检项目的众数与平均值填充.

通过数据预处理的数据中存在不同的孕检项目具有不同的单位和量纲的情况,实验通过归一化处理得到统一量纲的数据,消除不同的量纲对于模型带来的影响^[7].归一化采用式(1)所示的计算方法,其中 x 表示当前特征值, x_{\max} 、 x_{\min} 表示当前特征值的最大值和最小值. \bar{x} 为标准化后的特征值,标准化后的数据范围为 $[-1, 1]$.

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

在训练数据集中,孕妇孕检次数范围为1至84,为方便模型处理,需要对孕检数据进行标准化.标准化依据医学上对孕期的划分以及模型适合的数据格式进行处理.在医学上将孕期划分为孕前期、孕中期、孕后期3个阶段^[8],通常情况下孕妇的孕检频率依次增加,同时孕检结果的重要性也认为是依次增加.孕妇通常在孕早期内孕检次数较少,故本文将孕早期即孕周1至孕周12内的孕检记录为一次;孕中期即孕13周至

孕 28 周数量相对也较少, 分为 3 次, 即 [13-18] 周, [19-23] 周, [24-28] 周; 孕后期即孕 29 周至孕 40 周孕检次数较多, 每两周作为一个周期. 由于存在晚产的情况, 40 周以后还会出现孕检信息, 本文将 41 周之后作为一个新的孕检周期. 将同一孕检周期内的多次孕检结果的均值作为本次孕检周期内的孕检结果值, 若孕妇在某周期内无孕检数据, 则将该周期的孕检结果值均置为空值. 孕检数据的标准化如表 1 所示.

表 1 历次孕检数据划分

孕检周期	孕周	含义
孕期 1	[1-12]	1~12 周内孕检结果均值
孕期 2	[13-18]	13~18 周内孕检结果均值
孕期 3	[19-23]	19~23 周内孕检结果均值
孕期 4	[24-28]	24~28 周内孕检结果均值
孕期 5/6/7/8/9/10	[29-30], [31-32], ..., [39-40]	29~40 周内每两周结果作均值
孕期 11	[41-生产]	41 周及之后的结果作均值

1.2 特征构建

电子病历数据中不同的检查项目反映了孕妇不同层面的情况. 其中, 年龄、血型、孕次、病史等孕妇的

基本信息, 不受本次怀孕的影响, 反映出孕妇自身身体的基本素质, 因此将孕妇基础信息构建为孕妇基础特征群; 而首次孕检时记录的孕妇怀孕初始时的血压、体重、高危因素等, 反映出孕初期的基本情况, 可以据此评估孕妇在孕程初期的基本状态, 因此将首次孕检的数据构建为首次孕检特征群; 历次孕检结果记录了整个孕期的动态发展过程, 且具有时序的特点, 将其构建为历次孕检特征群. 而为了增加实验中模型对时序数据变化关系的学习能力, 本文将历次孕检的变化记录提取为新的特征群, 称为历次孕检结果变化特征群. 通过数据预处理及特征构建, 本文共构建了如表 2 所示的 4 个特征群. 其中, 历次孕检结果变化特征群记录处理后的不同孕检周期的结果变化值, 实验中孕期分为 11 段, 因此孕检结果变化共有 10 次记录, 若孕妇在该孕检周期内无孕检结果则将该差值置为空值, 若下一孕检周期有记录则下一周期差值为与最近的前次孕检周期结果的差值, 例如某孕妇仅进行两次孕检, 在孕期 1 内进行一次孕检, 孕期 10 进行一次孕检, 则该孕妇仅有第 9 次孕检结果变化有记录, 其余置为空值.

表 2 实验特征说明

特征群名	包含特征	说明	适用模型
孕妇基础特征群	血型 孕次 ...	包含孕妇自身的基本信息	LSTM/XGBOOST
首次孕检特征群	人工流产次数 血压 体质指数 ...	首次孕检结果	LSTM/XGBOOST
历次孕检特征群	高危因素 检查孕周 宫高 ...	除首次孕检外的其他孕检结果	LSTM/XGBOOST
历次孕检结果变化特征群	腹围 孕周差值 宫高变化 ... 腹围变化	历次孕检结果的变化	XGBOOST

2 构建基于 LSTM 和 XGBoost 的混合模型

孕妇产后出血受到多方面因素的影响, 实验数据同时包含时序数据和非时序数据两部分, 如上一章所介绍, 电子病历中记录的数据包括: 孕妇的身高、血型、孕次等基础信息, 为非时序数据; 孕妇历次孕检的数据, 如体重、血糖、腹围等, 此为时序数据. 模型需

要具有处理时序数据之间的变化关系, 同时, 在医疗应用中模型应具有较好的可解释性, 从而帮助医生根据经验对预测结果进行判断, 另外, 方便医生了解疾病的影响因素. LSTM 模型在时间序列数据预测具有良好的效果, 而 XGBoost 模型作为传统的集成模型在拥有良好性能的同时具有较强的可解释性, 在实际的疾病

预测应用中,两者都存在各自的缺陷,LSTM模型作为深度模型在拥有较好性能但难以对模型的结果做出解释,而XGBoost作为传统集成模型无法提取时序数据中的时序信息.本文利用LSTM模型提取数据中的时序信息结合XGBoost模型实现在模型拥有良好性能的同时对结果具有一定的可解释性.混合模型能够为医生的诊断提供参考意见,获取高危风险因素,方便针对性的对某些特征进行检查.

混合模型的构建分成两个部分.首先,构建LSTM模型数据集,利用LSTM模型对数据中的时序特征进行处理,得到预测结果;然后,利用LSTM模型结果构建新的特征,结合原有数据集组成新的数据集训练XGBoost模型,并利用训练后的模型得到最终结果.LSTM模型作为深度学习模型和集成模型的XGBoost在结构上有较大的差异,因此,混合模型能够较好地降低模型过拟合的风险.构建混合模型的流程如图1所示.

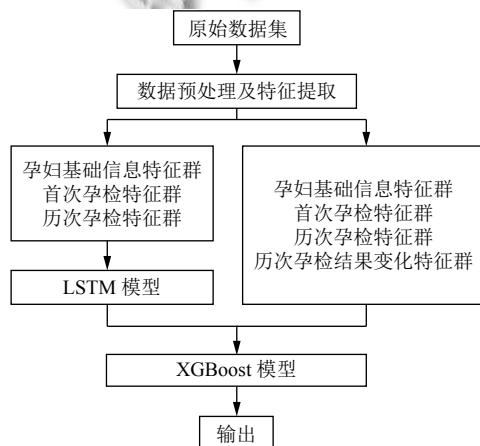


图1 模型流程图

2.1 LSTM模型

长短期记忆网络(Long Short Term Memory),简称LSTM,是循环神经网络(RNN)^[9]的优秀改进,属于反馈神经网络.相对于传统RNN网络,LSTM能够改善“梯度消失”的问题^[10],同时,LSTM在学习长期依赖的^[11]问题上有更加优秀的表现.LSTM模型的结构如图2所示.

LSTM在RNN结构的基础上增加每层结构的控制门,控制门分3类:Forget Gate(遗忘门)、Input Gate(输入门)、Output Gate(输出门)^[12].通过控制门的开关判断网络的状态在该层的输出是否达到设定的阈值,从而选择是否将其加入本层的计算中.

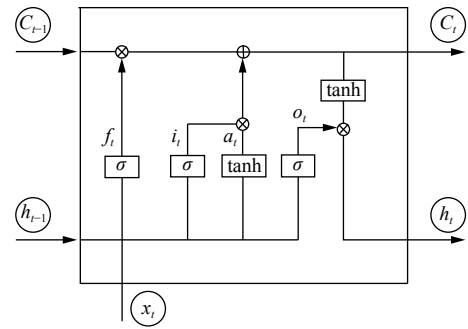


图2 LSTM结构

遗忘门在LSTM中控制神经元以一定的概率丢弃信息,根据上一序列隐藏状态 $h(t-1)$ 及本次序列的输入,利用激活函数得到门的输出 f_t ,如式(2)所示:

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (2)$$

其中, σ 为激活函数, W_f 为隐藏层到门的权重, U_f 为输入层到门的权重, b_f 为偏执向量.

输入门负责处理输入信息,输入门利用Sigmoid和tanh两个激活函数的输出的乘积更新神经元的信息,如式(3)、式(4)所示:

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (3)$$

$$a_t = \tanh(W_a h_{t-1} + U_a x_t + b_a) \quad (4)$$

更新神经元的信息来自两个部分,一是 C_{t-1} 与遗忘门的乘积,二是 i_t 和 a_t 的乘积,如式(5)所示:

$$C_t = C_{t-1} \cdot f_t + i_t \cdot a_t \quad (5)$$

在更新神经元后通过输出门控制得到 t 时刻神经元的输出 h_t ,如式(6)所示:

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

2.2 XGBoost模型

XGBoost是2014年被提出的分布式梯度提升算法(Boosting)工具库,其具有高效、灵活、鲁棒的特点,相对于其他Boosting方法^[13],XGBoost增加正则化项,能够防止模型过拟合.XGBoost目标函数如式(7)所示:

$$L(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (7)$$

其中, $\Omega(f_k)$ 如式(8)所示:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (8)$$

包含正则化项, \hat{y}_i 为模型的输出, y_i 为真实标签, f_k 为表示第 k 个基分类器, T 表示叶子数, ω 表示叶子

节点的权重, γ 为惩罚项.

利用泰勒公式^[14]将式 (7) 展开得到式 (9):

$$L^{(i)} = \sum_{i=1}^n \left[g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_i) \quad (9)$$

式中, g_i 与 h_i 以及 $\Omega(f_i)$ 表达式分别如式 (10)~式 (12) 如下:

$$g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \quad (10)$$

$$h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)}) \quad (11)$$

$$\Omega(f_i) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (12)$$

将式 (10)~式 (12) 代入式 (9) 得到叶子节点权重公式如式 (13) 所示:

$$\omega_j^* = -\frac{G_j}{H_j + \lambda} \quad (13)$$

式中, G_j 、 H_j 分别表示一阶梯度和二阶梯度在叶子节点 i 的值. 将式 (13) 带入式 (9) 得到目标函数的最优解如式 (14) 所示:

$$\hat{L}^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (14)$$

3 实验与分析

3.1 实验过程

本文在我国东部某医院 2018 年间产科所接诊的所有孕产妇的电子病历中, 选取了包含产后出血记录的 14 409 条样本构建数据集. 根据医学标准将胎儿娩出后 24 小时内产妇出血量超过 500 ml 的样本作为产后大出血的正样本. 其中正样本 701 条, 负样本 13 708 条. 随机抽取 70% 的正样本和 70% 的负样本组成训练集, 剩余 30% 的正样本和 30% 的负样本组成测试集. 在实验中采用 10 折交叉验证的方式寻找模型最优参数.

将 LSTM 模型应用于产后大出血的预测, 神经网络和网络层数直接影响最终效果^[15]. 实验中采用输入数据特征维度作为输入层的神经元数量; 输出层对应一个神经元表示是否会发生产后大出血. LSTM 隐藏层^[16]的节点数通常利用式 (15) 确定.

$$N = \sqrt{m+n} + a \quad (15)$$

其中, N 为隐藏节点数量, m 、 n 为输入、输出节点数, a 为区间 [1,10] 内的常数. 实验中输入节点个数为

59 个, 输出节点个数为 1 个, 根据式 (11) 得到隐藏节点数量为 8~18 个. 实验中通过利用同一数据集选择不同隐藏层节点的方法最终选择隐藏层节点数为 15 个.

将 LSTM 模型得到的结果作为新的特征加入构建新的数据集, 并利用该数据集训练 XGBoost 模型.

3.2 评价指标选取

本文目标为分类, 选取 ROC 曲线^[17]及曲线下面积 AUC 作为模型性能的评价标准. ROC 曲线的横标为假阳性率, 表示被划分为正类的结果中负实例占所有负实例的比例; 纵标为真阳性率, 表示被划分为正类的结果中正实例占所有正实例的比例. 根据选取的不同划分正负类的阈值, 形成 ROC 曲线. AUC 是一个概率值, 计算方法是 ROC 曲线下的面积, 取值范围为区间 [0.1,1], AUC 值越大表明该模型分类性能越好.

3.3 结果分析

XGBoost 模型在迭代 650 轮时取得最优效果, AUC 值为 0.70, ROC 曲线如图 3 所示.

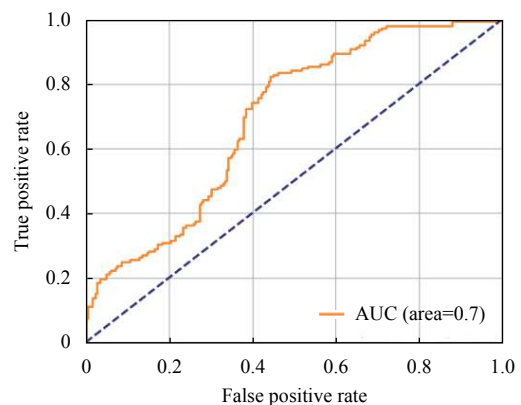


图3 预测产后出血的 ROC 曲线

从图 3 中可知, 基于 LSTM 和 XGBoost 的混合模型在利用孕妇数据和孕检数据对产后大出血进行预测能够在 40% 的假阳率下准确预测出 77% 的大出血数据, 能够为产前准备工作提供参考意见. 同时, 即使在本文中被误判为会发生产后大出血的孕妇, 相对于其他孕妇, 依然面临更高的产后大出血风险. 因此, 针对这些孕妇进行相关准备对进一步降低生产风险仍有重要意义.

XGBoost 模型给出的重要特征如图 4 所示. 通过输出特征重要程度, 可以判断哪些因素对产后大出血有更显著的影响, 从而辅助医生分析孕妇的生产

情况,是否要提前做好产后大出血的相关准备工作. LSTM 模型构建的新特征是最终模型结果中最重要的特征,可见 LSTM 和 XGBoost 的混合模型对于提高预测的准确率是有效的. 模型给出的重要特征如图 4 所示,其中胎盘相关因素、血压相关因素、年龄等在医学实践中都是经验丰富的医生预判孕妇产后大出血的重要依据,与医学先验知识相符. 另外体重、贫血、肝功能、肾功能、畸形、流产次数等因素对产后出血的也有不可忽略的影响. 由于因素较多且不够直观,即使是经验丰富的医生也难以直接从冗长的病历中直接得出结论,对缺乏经验的医生而言难度更高,准确性也更难把握. 因此,医生可以结合本模型的预测结果,提前采取治疗措施或分娩备血方案,降低产后大出血的发生概率.

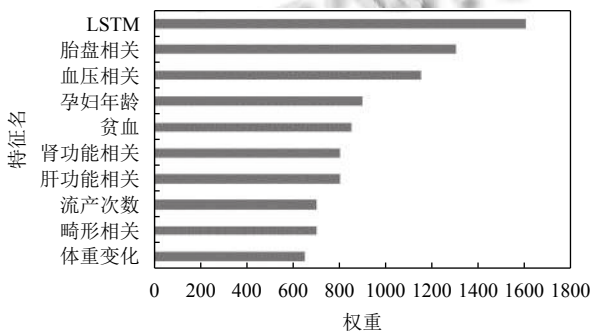


图 4 特征重要性

表 3 为各模型在同一测试集上的性能和运行时间,测试集包含 4000 条数据.

表 3 各模型 AUC 值及运行时间

模型	AUC 值	运行时间 (s)
LSTM	0.66	3.24
XGBoost	0.67	2.16
LSTM+XGBoost	0.70	6.03

结合图 4 和表 3 可看出 LSTM 模型构造的特征能够较大幅度提高模型的效果,而混合模型在三者中效果最好,同时运行时间也最长. 而 LSTM 模型在运行时间最长的前提下,单模型预测效果却最差,可能的原因是数据中正负样本偏差较大.

4 总结

LSTM 与 XGBoost 混合模型在构建数据集时应结合不同模型的特点. LSTM 模型对时间序列信息有较

强的处理能力,在数据集中将每次孕检数据的时间进行周期化处理,强化时间特征;针对 XGBoost 模型无法提取时间信息,需要提取孕检结果的变化,帮助模型获取隐藏的时序信息. 本文中,电子病历数据包含着孕妇孕期内的多类信息,不同类型的信息有不同的特点,为适应不同模型对数据的要求,需要对数据做出不同的处理,如孕妇多次孕检结果反映了孕期内孕妇身体状况和胎儿生长的动态变化,属于时间序列数据, LSTM 模型对时间序列信息有较强的处理能力,在数据集中将每次孕检数据的时间进行周期化处理,强化时间特征;针对 XGBoost 模型无法提取时间信息,需要提取孕检结果的变化,帮助模型获取隐藏的时序信息. 将 LSTM 模型对时间序列的处理能力和 XGBoost 模型的可解释性结合,对孕妇产后大出血作出预测. 实验结果表明,利用 LSTM 模型处理时间序列数据并将预测结果作为新特征训练 XGBoost 模型相对于原有单模型能够有效提高模型的性能,因此利用混合模型对孕妇产后大出血进行预测是可行的. XGBoost 给出的重要特征能够帮助医生筛选高危患者,通过提前备血等方式降低大出血可能带来的风险. 本文采用混合模型在性能提升的同时模型的运行时间长于单模型,如何降低模型的运行时间,满足大数据条件下的预测是后期工作的重点.

参考文献

- 王杨, 张薇, 席红, 等. 探讨妊娠期高血压疾病并发 HELLP 综合征的临床特点及母婴结局. 中国妇幼保健, 2017, 32(17): 4096-4099.
- 于国玲. 浅析产后大出血的病因和治疗. 内蒙古医学杂志, 2012, 44(13): 11-13.
- Liu L, Shen J, Zhang M, *et al.* Learning the joint representation of heterogeneous temporal events for clinical endpoint prediction. Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, LA, USA. 2018. 109-116.
- Chokwijitkul T, Nguyen A, Hassanzadeh H, *et al.* Identifying risk factors for heart disease in electronic medical records: A deep learning approach. Proceedings of BioNLP 2018 workshop. Melbourne, Australia. 2018. 18-27.
- Chen TQ, Guestrin C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA. 2016. 785-794.

- 6 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- 7 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille, France. 2015. 448–456.
- 8 Salomon LJ, Alfirevic Z, Berghella V, *et al.* Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound in Obstetrics & Gynecology*, 2011, 37(1): 116–126.
- 9 Graves A. Supervised sequence labelling. Graves A. *Supervised Sequence Labelling with Recurrent Neural Networks*. Berlin, Heidelberg: Springer, 2012. 5–13.
- 10 LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324. [doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791)]
- 11 Zhang SL, Liu C, Jiang H, *et al.* Nonrecurrent neural structure for long-term dependence. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2017, 25(4): 871–884. [doi: [10.1109/TASLP.2017.2672398](https://doi.org/10.1109/TASLP.2017.2672398)]
- 12 张尧. 激活函数导向的 RNN 算法优化[硕士学位论文]. 杭州: 浙江大学, 2017.
- 13 Bühlmann P, Hothorn T. Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science*, 2007, 22(4): 477–505. [doi: [10.1214/07-STS242](https://doi.org/10.1214/07-STS242)]
- 14 Ismail MEH, Stanton D. Applications of q -Taylor theorems. *Journal of Computational and Applied Mathematics*, 2003, 153(1–2): 259–272.
- 15 Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. *Neural Computation*, 2000, 12(10): 2451–2471. [doi: [10.1162/089976600300015015](https://doi.org/10.1162/089976600300015015)]
- 16 Sankaran S, Audet C, Marsden AL. A method for stochastic constrained optimization using derivative-free surrogate pattern search and collocation. *Journal of Computational Physics*, 2010, 229(12): 4664–4682. [doi: [10.1016/j.jcp.2010.03.005](https://doi.org/10.1016/j.jcp.2010.03.005)]
- 17 沈小玲, 沈雅芬, 金建芬, 等. 应用 ROC 曲线评价心血管疾病住院病人护理评估模型的临床价值. *中国农村卫生事业管理*, 2012, 32(6): 577–579.