

基于深度森林的产前胎儿监护不平衡多分类判别*

郭 傲

陈妍获 魏 航 陈沁群 洪佳明

(1 广州中医药大学医学信息工程学院 广州 510006 (广州中医药大学医学信息工程学院 广州 510006)
2 广东工业大学计算机学院 广州 510006)

李琼娜

郝志峰

(1 广州东仁医院 广州 510442 (佛山科学技术学院数学与大数据学院 佛山 528000)
2 广州三瑞医疗器械有限公司 广州 510520)

〔摘要〕 采用深度森林框架构建基于不平衡电子胎心宫缩监护数据的多分类判别模型, 验证模型有效性, 结果表明该模型预测性能较好, 极大降低误判率, 在产前胎儿健康状况智能评估中有良好应用前景。

〔关键词〕 产前胎儿监护; 不平衡多分类; 深度森林; 胎心宫缩监护

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2021.03.009

Imbalanced Multi-classification Discrimination for Antenatal Fetal Monitoring Based on Deep Forest GUO Ao, 1School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou 510006, 2School of Computers, Guangdong University of Technology, Guangzhou 510006, China; CHEN Yandi, WEI Hang, CHEN Qinqun, HONG Jiaming, School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; LI Qionghua, 1Guangzhou Dongren Hospital, Guangzhou 510442, 2Guangzhou Sunray Medical Apparatus Co. Ltd., Guangzhou 510520, China; HAO Zhifeng, School of Mathematics and Big Data, Foshan University, Foshan 528000, China

〔Abstract〕 The multi-classification discrimination model based on unbalanced Cardiotocography (CTG) data is built by Deep Forest (DF) framework. The validity of the model is verified, and the results show that the prediction performance of the model is good, the misjudgment rate is greatly reduced, and the model has a good application prospect in the intelligent assessment of antenatal fetal health status.

〔Keywords〕 antenatal fetal monitoring; imbalanced multi-classification; Deep Forest (DF); Cardiotocography (CTG)

〔收稿日期〕 2020-06-24

〔作者简介〕 郭傲, 硕士研究生, 发表论文 1 篇; 通讯作者: 魏航, 讲师, 发表论文 31 篇。

〔基金项目〕 国家自然科学基金资助项目“高阶网络模体聚类算法与应用研究”(项目编号: 61976052); 广东省医学科研基金资助项目“基于不平衡 CTG 数据的产前智能胎儿监护评价模型的研究”(项目编号: A2019428)。

1 引言

1.1 定义

胎心宫缩监护 (Cardiotocography, CTG) 记录胎心率变化及其与宫缩之间的时间关系, 其目的是确定哪些胎儿可能缺氧, 以指导对胎儿健康的进一步评估, 确定是否需要孕妇进行剖腹产或辅助阴道分娩^[1]。目前 CTG 在胎儿监测方面已被广泛应用, 其对社会、国家、孕产妇家庭和医院都具有非常重要的意义。

1.2 研究背景

产前胎儿监护指南^[2]将 CTG 所得的胎儿健康状况评估为正常、可疑和异常 3 类。相关文献显示胎儿窘迫发病率约为 3% ~ 39%^[3], 异常类案例相对正常类和可疑类较少。现有大部分基于机器学习的胎儿监护模型研究是基于样本分布平衡和分类精度最大化的算法, 往往将分布不平衡情况归为正态分布, 忽略 CTG 数据不平衡问题^[4-11], 所得可疑和异常病例准确率大多在 45% ~ 82% 和 66% ~ 94% 之间^[6-11], 无法应用于产前胎儿监护。另外无应激试验 (Non - Stress Test, NST) 诊断标准中正常类和可疑类判别标准相互交叉^[6], 易造成正常类和可疑类相互之间误判, 若将可疑类误判为正常类则很有可能耽误胎儿检查与治疗, 对胎儿和孕妇健康带来极大风险。一些研究^[4-5]为解决正常类和可疑类判读困难问题, 直接删去可疑案例, 仅研究正常类和异常类, 将多分类转化为二分类, 这种处理使得机器学习模型结果灵敏度较低, 不能满足胎儿监护实际需要。

1.3 研究目的

鉴于此, 本文提出利用深度森林框架构建基于不平衡 CTG 数据的多分类判别模型。通过深度森林多粒度扫描获取样本数据增强特征^[12], 避免不同标签之间特征的过渡交叉, 从而有效地解决正常类和可疑类判读标准交叉问题, 在此基础上利用级联森林将多个树形结构模型进行多次深度迭代集成学习, 得到一个最好的预测效果模型。

2 实验数据

2.1 数据来源

本研究从公开可用的加州大学欧文分校机器学习知识库 (University of California Irvine Machine Learning Repository, UCI)^[13]选择产前胎心宫缩监护数据集 (网址: <https://archive.ics.uci.edu/ml/datasets/Cardiotocography>), 该数据集包含 21 个胎心宫缩特征和 3 位产科专家对胎儿状态“正常、可疑和异常”判读一致的结果。CTG 数据集共有数据 2 126 例, 其中胎儿状态为正常类、可疑类、异常类的分别有 1 655 例、295 例和 176 例。

2.2 数据探索

胎儿状态评估结果数量对比, 见图 1。

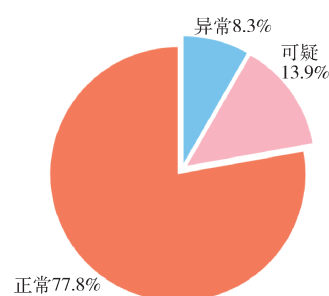


图1 胎儿状态评估结果对比

可见正常类数量远超前于可疑类和异常类, 正常类数量很多, 可疑类和异常类数量极少, 3 类数据分布极不平衡。利用主成分分析 (Principal Component Analysis, PCA) 方法对数据特征进行探索, 研究胎心宫缩特征和胎儿状态分布关系。主成分降维分布, 见图 2。可疑案例和正常案例有大面积交叉, 由此可知正常类和可疑类之间较易发生判读失误。

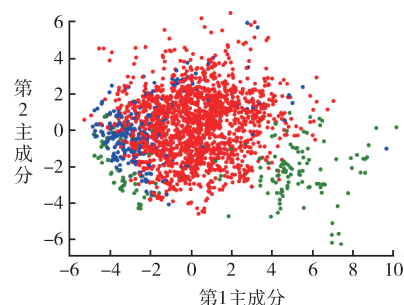


图2 主成分分析降维分布

注: 红色点代表正常案例, 蓝色点代表可疑案例, 绿色点代表异常案例。

3 实验方法

3.1 深度森林算法

3.1.1 概述 深度森林（Deep Forest，DF）是在深度学习理论以及神经网络启发下，以随机森林算法（Random Forest，RF）为基础的一种有监督机器学习算法。利用多粒度扫描对原始输入特征进行转换以增强特征表达能力，再通过级联森林做逐层表征学习。

3.1.2 多粒度扫描 CTG 数据中正常类和可疑类在评判标准特征范围的相互交叉导致预测算法对这两类预测准确度低于异常类，因此找出特征之间的关系对提升预测算法准确度至关重要。为提高深度森林算法预测效果以及尽可能地挖掘序列数据特征顺序关系，通过深度森林多粒度扫描对样本特征进行提取。对能有效地找出各个特征之间顺序关系的预测算法，其预测准确率可得到明显提升^[13]。

3.1.3 级联森林 深度神经网络中的表征学习主要依赖于对原始特征的逐层处理，基于深度神经网络的启发，深度森林算法利用级联森林体现算法深度学习过程。深度森林深度可自动由算法确定而不需要人为设置参数。当级联森林不再增长时，深度

森林将每个森林预测类别对应结果相加后取平均值，类别概率最大值即为该样本最终预测结果，计算公式如下：

$$label = \underset{i \in \{1,2,3\}}{\operatorname{argmax}} \frac{\sum_{t=1}^4 y_{t,i}}{4}$$

(1)

3.2 基于深度森林的产前胎儿监护判别模型

3.2.1 深度森林网络构建 基于深度森林的胎儿监护预测模型整体流程，见图 3。深度森林的输入向量为 21 维的 CTG 样本数据，多粒度扫描阶段的滑动窗口分别为 2 维、3 维和 6 维，在经过随机森林和完全随机森林（Completely Random Forest，CRF）^[14]训练后分别产生 120 维、114 维、96 维的特征向量，将多粒度扫描生成的首个特征向量作为级联森林的第 1_A 级输入，对于胎儿监护 3 分类问题，在分别经过 4 种森林预测分类后得到 4 个 3 维的类向量，共 12 维的类向量与首层输入的特征向量合并拼接为 132 维的级联森林第 1_B 级的输入向量，而第 1_B 层级的输出会与多粒度扫描生成的第 2 个特征向量合并拼接为下一层级的输入，经过多层迭代，直到深度森林算法验证整体模型准确率不再提高时，模型训练结束。

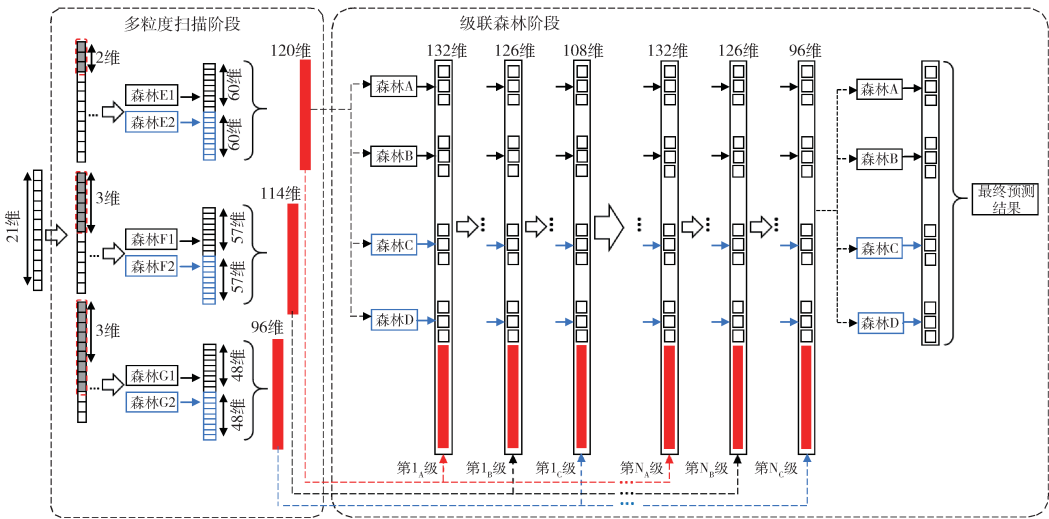


图3 深度森林算法流程

3.2.2 实验参数设置 经过多次试验，确定深度森林的级联森林阶段 4 种模型分别为随机森林、完

全随机森林、加权随机森林（Weighted Random Forest，WRF）和梯度提升决策树（Gradient Boosting

Decision Tree, GBDT)^[15]，每个森林模型都由 200 个决策树组成；多粒度扫描阶段选用完全随机森林和随机森林模型，每个模型包含 500 个决策树。通过对数据进行五折交叉验证来确定最佳模型深度，实验表明当深度森林层数为 5 时，模型具有最高分类精度。

4 实验结果

4.1 概述

基于 CTG 数据，采用五折交叉验证方法来获得模型评估指标，为评估深度森林算法对不平衡胎儿监护数据的预测能力，选取随机森林、梯度提升决策树、支持向量机（Support Victor Machine, SVM）^[16]、K 最邻近分类（K - Nearest Neighbor, KNN）^[17]和朴素贝叶斯（Naive Bayes, NB）^[18]5 种传统机器学习算法和深度神经网络（Deep Neural Network, DNN）^[19]进行预测，将算法预测结果与深度森林算法预测结果进行对比分析，以验证深度森林模型有效性。预测算法程序机器配置 CPU 为 Core i7 - 8550U，1.99GHz，内存 8GB，系统为 Win-

dows10 家庭中文版 64 位，编程环境为 Python 3. 6. 5 及 PyCharm Professional 2018. 1. 4。

4.2 模型评估指标

为验证所提出的深度森林模型在不平衡 CTG 数据下的性能，采用准确率（Accuracy）、精确率（Precision）、召回率（Recall，也称敏感度（Sensitivity））、综合评价指标（F1 Score）以及观测者操作特性曲线（Receiver Operating Characteristic Curve, ROC）^[20]曲线下面积值（The Area Under the Curve, AUC）指标，计算公式如下：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
 (2)

$$Precision = \frac{TP}{TP + FP}$$
 (3)

$$Recall(Sensitivity) = \frac{TP}{TP + FN}$$
 (4)

$$F1 = \frac{2}{1/Precision + 1/Recall}$$
 (5)

4.3 结果与分析

4.3.1 与传统机器学习模型和深度神经网络的对比（表 1）

表 1 算法总体分类评价对比

评价指标		精确率	召回率（灵敏度）	F1	平均 F1	AUC	平均 AUC
深度森林	正常	0. 970 0	0. 970 0	0. 970 0	0. 920 1	0. 987 8	0. 989 5
	可疑	0. 871 0	0. 843 8	0. 857 1	-	0. 982 5	-
	异常	0. 903 2	0. 965 5	0. 933 3	-	0. 998 3	-
NB	正常	0. 912 1	0. 903 9	0. 908 0	0. 651 5	0. 915 9	0. 877 5
	可疑	0. 514 3	0. 562 5	0. 537 3	-	0. 870 7	-
	异常	0. 538 5	0. 482 8	0. 509 1	-	0. 845 9	-
SVM	正常	0. 953 4	0. 921 9	0. 937 4	0. 838 1	0. 959 6	0. 964 5
	可疑	0. 641 0	0. 781 3	0. 704 2	-	0. 945 3	-
	异常	0. 923 1	0. 827 6	0. 872 7	-	0. 988 6	-
KNN	正常	0. 918 8	0. 952 0	0. 935 1	0. 803 0	0. 889 9	0. 892 0
	可疑	0. 678 0	0. 6250	0. 650 4	-	0. 860 2	-
	异常	0. 954 5	0. 724 1	0. 823 5	-	0. 926 0	-
RF	正常	0. 953 4	0. 982 0	0. 967 5	0. 895 4	0. 990 8	0. 989 1
	可疑	0. 886 8	0. 734 4	0. 803 4	-	0. 978 1	-
	异常	0. 900 0	0. 931 0	0. 915 3	-	0. 998 5	-

续表 1

GBDT	正常	0.958 7	0.976 0	0.967 3	0.915 2	0.983 2	0.985 1
	可疑	0.879 3	0.784 6	0.829 3	—	0.972 8	—
	异常	0.933 3	0.965 5	0.949 2	—	0.999 4	—
DNN	正常	0.921 7	0.987 6	0.953 5	0.832 1	0.924 6	0.914 1
	可疑	0.928 6	0.609 4	0.735 8	—	0.885 7	—
	异常	0.821 4	0.793 1	0.807 0	—	0.932 1	—

可以看出深度森林可以更有效地解决不平衡的胎心宫缩图数据问题：正常类的精确度，可疑类的灵敏度、F1 和 AUC，异常类的灵敏度以及整体模型平均 F1 和平均 AUC 值均为最高。与同样具有层级结构的 DNN 相比，深度森林可疑类召回率（灵敏度）提高 23.44%，过低的召回率会导致真实的可疑类被误判为正常类，这将造成延误治疗等严重

的误诊后果，深度森林可疑类的精确度较 DNN 低 5.76%，将少量正常类误判为可疑类仅会造成多一次对胎儿的检查。此外从部分预测指标结果与 GBDT 和 RF 相近可以看出深度森林很好地集成级联森林中基模型预测能力，获得更好的整体预测性能。各模型 ROC 曲线面积比较，见图 4。

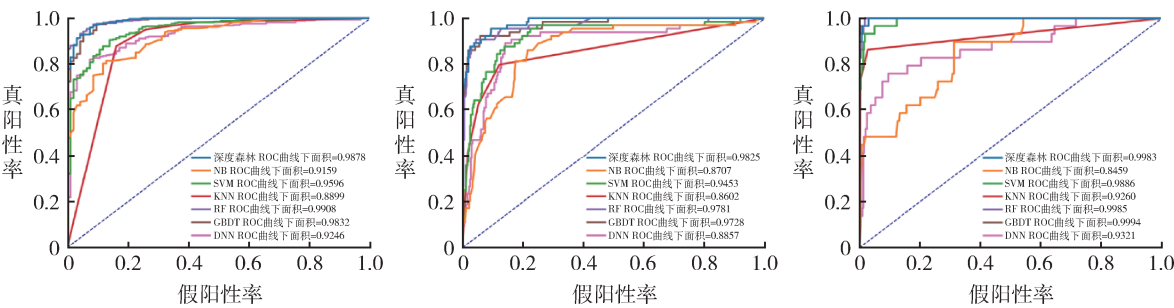


图 4 CTG 多分类 ROC 曲线

深度森林模型预测性能均优于传统机器学习模型和 DNN。模型和随机分类误判分类雷达图，见图 5。

6 种错误分类中最危险的情况是将异常类或可疑类误判为正常类，这将导致胎儿错过最佳治疗时间，危及胎儿健康。此外将正常状态误判为异常也十分危险，可能导致过度治疗以及对胎儿的过度干预，甚至造成不必要的剖腹产。结果表明与其他模型相比，深度森林模型错误分类率最低，且大大降低可疑类误判为正常类的错误分类率，显著提升对异常类识别能力。

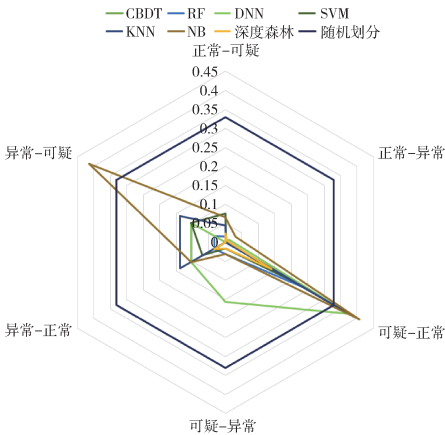


图 5 7 种分类器错分雷达

4.3.2 与国内外现有产前胎儿监护模型的对比（表 2）可以看出本文所提出的模型在异常和可疑类预测精度、平均 F1 和准确度方面均高于现有产前胎儿检测模型，平均 F1 评分较 DT-AdaBoost 模型提高 0.78%^[23]，表明深度森林在多分类 CTG 数据中判别正常和可疑类具有优势。

表 2 现有 CTG 判别模型总体精度和分类性能对比

模型	特征个数	正常类（%）	可疑类（%）	异常类（%）	平均 - F1（%）	准确度（%）
BP ^[19]	21	97.84	45.14	97.24	80.07	91.31
GRNN ^[21]	21	95.70	73.92	84.88	84.83	91.86
PNN ^[21]	21	95.91	73.81	85.45	85.06	92.14
MLPNN ^[21]	21	95.00	68.43	80.50	81.31	90.36
RF ^[22]	21	96.40	79.60	91.20	89.07	93.60
LAGA ^[22]	6	96.83	79.15	89.41	88.46	93.89
DT - AdaBoost ^[23]	21	97.15	83.69	92.84	91.23	95.01
DA ^[24]	10	89.69	58.50	65.58	71.26	82.03
LS - SVM - PSO - BDT ^[25]	21	96.02	72.98	79.18	82.73	91.58
DT ^[24]	10	93.31	60.09	66.43	73.28	86.31
深度森林（本文）	21	97.00	85.71	93.33	92.01	95.07

5 结语

本研究提出一种用于产前胎儿监护不平衡多分类的深度森林模型，该模型使用不平衡的 CTG 数据准确评估胎儿健康状况，基于多粒度扫描对样本特征进行提取，获取特征之间的关系，再利用级联森林进行多种模型的集成以及层次学习，以解决 CTG 数据不平衡以及判读标准相互交叉导致的分类错误问题，减少正常和可疑类别误判。本模型将人工智能与产前胎儿监护相结合，可用于新型的产前智能胎儿监护系统，孕妇通过可穿戴终端采集 CTG 信号，将数据远程实时发送至云服务器计算平台，以实现远程医疗，使偏远、贫困地区孕妇能享受到方便、及时、准确的胎儿监护服务。深度森林模型使用胎儿心率宫缩信号中提取的临床特征存在一定误差，在今后的研究中一方面对 CTG 特征提取进行改进，另一方面尝试直接处理 CTG 信号并进行建模，进一步提高模型分类准确率。

参考文献

1 AlfIREVIC Z, Devane D, Gyte G M L. Continuous Cardiotocography (CTG) as a Form of Electronic Fetal Monitoring (EFM) for Fetal Assessment during Labour [J]. Cochrane Database of Systematic Reviews (Online), 2006, 5 (3):

CD006066.

2 Xie Xing, Yan Wenli. Obstetrics and Gynecology [M]. Beijing: People's Medical Publishing House, 2013: 118 - 120.

3 张蓓, 闫辉波. 孕妇血雌三醇及脐动脉血流 S/D 比值检查对胎儿窘迫诊断的价值 [J]. 血栓与止血学, 2019, 25 (1): 23 - 26.

4 Kouskouti C, Regner K, Knabl J, et al. Cardiotocography and the Evolution into Computerised Cardiotocography in the Management of Intrauterine Growth Restriction [J]. Archives of Gynecology and Obstetrics, 2017, 295 (4): 811 - 816.

5 谢幸, 孔北华, 段涛. 妇产科学 [M]. 北京: 人民卫生出版社, 2018: 54 - 56

6 程志厚, 宋树良. 胎儿电子监护学 [M]. 北京: 人民卫生出版社, 2001: 13 - 114.

7 中华医学会围产医学分会. 电子胎心监护应用专家共识 [J]. 中华围产医学杂志, 2015 (7): 486 - 490.

8 Haran S S, Everett T R. Antenatal Fetal Wellbeing [J]. Obstetrics, Gynaecology and Reproductive Medicine, 2017, 27 (2): 44 - 49.

9 Grivell R M, AlfIREVIC Z, Gyte G M, et al. Antenatal Cardiotocography for Fetal Assessment Cochrane Database [J]. Cochrane Database of Systematic Reviews (Online), 2010, 12 (1): CD007863.

10 田才宣. NST 评分法和改良 Fischer 评分法评价胎监结果与新生儿出生情况相关性 [J]. 中国当代医药, 2009,

- 16 (22): 28 – 29.
- 11 Santo S, Ayres – de – Campos D, Costa – Santos C, et al. Agreement and Accuracy Using the FIGO, ACOG and NICE Cardiotocography Interpretation Guidelines [J]. *Acta Obstetrica et Gynecologica Scandinavica*, 2017, 96 (2): 166 – 175.
- 12 Zhou ZH, Feng J. Deep Forest: towards an alternative to deep neural networks [EB/OL]. [2018 – 01 – 31]. <https://arxiv.org/abs/1702.08835>
- 13 Cho K, Van Merriënboer B, Gulcehre C, et al. Learning Phrase Representations Using RNN Encoder – decoder for Statistical Machine Translation [C]. Doha: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2014: 1724 – 1734.
- 14 王鑫, 汪晋宽, 刘志刚, 等. 基于随机森林的认知网络频谱感知算法 [J]. *仪器仪表学报*, 2013 (11): 73 – 79.
- 15 Ke Guolin, Meng Qi, Thomas F. LightGBM: a highly efficient gradient boosting decision tree [C]. Long Beach: Proc of the 31st Conference on Neural Information Processing Systems, 2017.
- 16 Platt J C. Sequential Minimal Optimization: a fast algorithm for training support vector machines [EB/OL]. [2020 – 07 – 30]. https://www.researchgate.net/publication/2624239_Sequential_Minimal_Optimization_A_Fast_Algorithm_for_Training_Support_Vector_Machines
- 17 Mangalova E, Shesterneva O. K – nearest Neighbors for GEF-Com2014 Probabilistic Wind Power Forecasting [J]. *International Journal of Forecasting*, 2016, 32 (3): 1067 – 1073.
- 18 John G H, Langley P. Estimating Continuous Distributions in Bayesian Classifiers [C]. San Mateo: Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI – 1995), 1995: 338 – 345.
- 19 Sundar C, Chitradevi M C M, Geetharamani G. Classification of Cardiotocogram Data Using Neural Network Based Machine Learning Technique [J]. *International Journal of Computer Applications*, 2013, 47 (14): 19 – 25.
- 20 Landgrebe T C W, Duin R P W. Efficient Multiclass ROC Approximation by Decomposition via Confusion Matrix Perturbation Analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30 (5): 810 – 822.
- 21 Yılmaz E. Fetal State Assessment from Cardiotocogram Data Using Artificial Neural Networks [J]. *Journal of Medical and Biological Engineering*, 2016, 36 (6): 820 – 832.
- 22 Sindhu R, Bahari J A, Hariharan M, et al. A Novel Clinical Decision Support System Using Improved Adaptive Genetic Algorithm for the Assessment of Fetal Well – Being [J]. *Computational and Mathematical Methods in Medicine*, 2015 (2015): 1 – 11.
- 23 Karabulut E M, İbrikci T. Analysis of Cardiotocogram Data for Fetal Distress Determination by Decision Tree Based Adaptive Boosting Approach [J]. *Journal of Computer and Communications*, 2014, 2 (9): 32.
- 24 Huang M L, Hsu Y Y. Fetal Distress Prediction Using Discriminant Analysis, Decision Tree, and Artificial Neural Network [J]. *Journal of Biomedical Science and Engineering*, 2012, 5 (9): 526.
- 25 Yılmaz E, Kılıkçer Ç. Determination of Fetal State from Cardiotocogram Using LS – SVM with Particle Swarm Optimization and Binary Decision Tree [J]. *Computational & Mathematical Methods in Medicine*, 2013 (2013): 487179.

《医学信息学杂志》开通微信公众号

《医学信息学杂志》微信公众号现已开通, 作者可通过该平台查阅稿件状态; 读者可浏览当期最新内容、过刊等; 同时提供国内外最新医学信息研究动态、发展前沿等, 搭建编者、作者、读者之间沟通、交流的平台。可在微信添加中找到公众号, 输入“医学信息学杂志”进行确认, 也可扫描右侧二维码添加, 敬请关注!



《医学信息学杂志》编辑部