

# Prediction of Delivery Mode from Cardiotocography and Electronic Medical Records Using Machine learning

Xue Kang<sup>1</sup>, Rongdan Zeng<sup>1</sup>, Hao Yi<sup>1</sup>, Chuan Wang<sup>1</sup>, Mujun Liu<sup>2</sup>, Zheng Zheng<sup>3</sup>,  
Yaosheng Lu<sup>1,4</sup>, Huijin Wang<sup>1</sup> and Jieyun Bai<sup>1,4\*</sup>

<sup>1</sup> College of Information Science and Technology, Jinan University, Guangzhou, China;

<sup>1</sup> College of Science & Engineering, Jinan University, Guangzhou, China;

<sup>3</sup> Department of Obstetrics, Preterm Birth Prevention and Treatment Research Unit, Guangzhou Women & Children's Medical Center, Guangzhou Medical University, Guangzhou, China

<sup>4</sup> Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Information Technology, Jinan University, Guangzhou 510632, China

\* Correspondence: bai\_jieyun@126.com; Tel.: +86 13274618376 (J.B.)

**Background:** Making effective decisions based on CTG data is a challenge. The American College of Obstetrics and Gynecology suggests that obstetricians in the proposal of delivery type should take into account the maternity's prenatal electronic medical information and risk factors such as pregnancy complications.

**Method:** This paper was conducted in two parts to identify the delivery type: First, traditional CTG features and meaningful Electronic Medical Record (EMR) features were extracted from the available CTG and prenatal medical records respectively. Then, several prediction models (including Logistic Regression (LR), Naive Bayes (NB), K-Nearest-Neighbor (KNN), Random Forest (RF), Decision Tree (DT), AdaBoost, XGBoost, Gradient Boosting Decision Tree (GBDT) and Bagging) were trained to classify patients according to their delivery type: vaginal and cesarean delivery. An association-based feature selection algorithm was employed on the features sets to remove the unnecessary features to improve the performance of classifiers. The data sets were composed of 783 signals collected between 2017 and 2020.

**Result:** The prediction models based on CTG features and EMR features were in terms of accuracy, sensitivity, specificity, and area under the curve receiver operating characteristic (AUC-ROC) in the outcomes. LR achieved the highest specificity (55.33%), NB achieved the highest level of accuracy (65.57%) and sensitivity (85.9%), AdaBoost achieved the highest AUC-ROC (57.91%). As expected, the specificity of the ensemble-learning classifier was 9.43% higher than the basic classifier, the specificity of all classifiers was improved by 4.5% after introducing the EMR features.

**Conclusion:** According to the obtained results, machine learning was proved with high value when predicting cesarean delivery, and was a useful tool for reducing the false-positive rate and unnecessary operative interventions by employing CTG data and EMR information.