EasyChair Preprint

№ 1722

# SEM-based study for interpretability of intelligent prenatal fetal monitoring models

Li-Ting Huang, Rui-Chu Cai, Zhi-Ying Jiang, Li Li,
Qin-Qun Chen, Jia-Ming Hong, Zhi-Feng Hao and Hang Wei

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 20, 2019

**Noname manuscript No.**
(will be inserted by the editor)

# SEM-based study for interpretability of intelligent prenatal fetal monitoring models

**Li-ting Huang[1,2] · Rui-chu Cai[1] · Zhi-ying Jiang[2] · Li Li[3,5] · Qin-qun Chen[2] · Jia-ming Hong[2] · Zhi-feng Hao[4] · Hang Wei[2]***
**[1]School of Computers, Guangdong University of Technology, China**
**[2]School of Medical Information Engineering, Guangzhou University of Chinese Medicine, China**
**[3]Tianhe District People's Hospital, First Affiliated Hospital of Jinan University, China**
**[4]School of Mathematics and Big Data, Foshan University, China**
**[5]Guangzhou Sanrui Medical Devices Co., LtdChina**
***Corresponding author : Hang Wei, crwei@gzucm.edu.cn**

Received: date / Accepted: date

**Abstract** Widely available computerized cardiotocography (CTG) data and machine learning methods present an opportunity to improve the accuracy and scalability of automated CTG analysis for prenatal fetal monitoring. However, their interpretabilities are not clear enough, due to superficially dependent on data. In this paper, we present a study on an interpretable reference for machine learning-based prenatal fetal monitoring models. CTG characteristics are classified into baseline category, variability category, acceleration category, deceleration category and uterine contraction (UC) category via a measurement model. Then structural equation models (SEMs) are introduced to derive the causal relationship between the CTG categories and explore their impact on the fetus status. The experimental results show that the variability category predicts the baseline category and UC category has a predictive effect on the deceleration category. In addition, the outcomes, that variability category and the acceleration category have a greater impact on the identification of fetal status while the baseline category has less impact, explain the prenatal fetal monitoring model based on weighted random forests. In summary, our study validate prenatal fetal monitoring clinical knowledge and provide an interpretable reference for intelligent prenatal fetal monitoring models.

**Keywords** : Prenatal fetal monitoring , Cardiotocography , Structural equation model , Interpretability , Random forests

Address(es) of author(s) should be given

## 1 Introduction

Cardiotocography (CTG) is important in fetal health monitoring. It helps early diagnosis of abnormalities, such as congenital heart defects, fetal distress or hypoxia, by responding to which obstetricians can take measures to prevent irreversible damage to the fetus [1]. Non-stress testing (NST) was widely introduced into prenatal fetal monitoring in the late 1960s and is still widely used due to its low cost, ease of operation, and non-invasiveness [2]. In NST, the interpretation of CTG, including fetal heart rate and uterine contraction signals, has so far largely depended on obstetricians. However, the demand for fetal monitoring has increased dramatically, while the shortage of specialist obstetricians is still severe, and the level of fetal monitoring in primary or rural hospitals is quite low [3]. Therefore, many researchers establish prenatal fetal monitoring models, through the use of uterine contractions (UC) and fetal heart rate ( FHR). In general, the current machine learning research methods for prenatal CTG features interpretation are mainly classification algorithms, lifting algorithms and hybrid algorithms. These methods achieved good experimental results, which is a high accuracy rate. However, the "accuracy rate" only illustrates that the method is feasible in the knowledge category. For the algorithm, transparency mainly reflects the interpretability of the model, that is, whether humans can understand or explain the conclusions drawn by the machine[4]. The reasoning process that cannot be explained artificially is likely to be meaningless. It is even more difficult to accept in the medical field for patients, due to the disease has complex risk correlation and is vital to life, and must be very rigorous in development[5].

Strengthening the interpretability of the problem world in machine learning. On the one hand, it allows users to better understand the decision-making process of the machine learning system, which is conducive to adding people's trust in the model. On the other hand, it can provide users with an operable interaction mode, which enables people's experience to intervene in data-driven modeling and decision making, to realize traceability, supervision, guidance, and correction of the analysis decision process, thereby improving system performance and performance[6].

Based on the above-mentioned medical AI and strict medical reality, we choose Sisporto data to use SEM model to achieve the following goals: 1) explore the causal relationship between hidden variables; 2) explore the causal relationship of hidden variables to fetal status; 3) use Causality validates clinical knowledge and interprets the results of the best machine learning model, the result of weighted random forests.The experimental results show that the variant category predicts the baseline category and that the contraction has a predictive effect on the deceleration category, validating the SisPorto 2.0[7] (CTG analysis program) and prenatal fetal monitoring clinical knowledge, respectively. In addition, the variability category and the acceleration category have a greater impact on the identification of fetal status, while the baseline category has less impact, explaining the importance distribution of CTG forests.

## 2 Related works

### 2.1 CTG and its interpretations in antenatal fetal monitoring

During antenatal fetal monitoring, it is necessary to record cardiotocography (CTG), including fetal heart rate (FHR) and uterine contraction (UC) [8−9]. Obstetricians usually assess fetal health status based on the characteristics of CTG, such as baseline, variation, deceleration, and uterine contraction. There are two existing evaluation methods for antenatal CTG, namely "scoring methods" and "ranking methods".

As for the scoring methods, there have NST, Krebs, Fischer, and modified Fischer, but these scoring methods have some shortcomings. Firstly, the scoring systems are not uniform, ranging from 10 to 12. Secondly, the scoring items are inconsistent [9−11]. Moreover, these methods cannot directly define whether the status of the fetus is normal.

In the ranking methods for antenatal CTG, differences exist in different countries. The mainstream antenatal fetal monitoring guidelines include Canadian SOGC [12], American ACOG, British NICE, International Federation of Obstetrics and Gynecology FIGO [13−14] and Chinese expert consensus. SOGC, ACOG, NICE and FIGO all adopt three-level evaluation among them. But in Chinese experts' consensus, the conditions of fetuses are only divided into two levels, namely as "reactive" and "non-response"[15]. Besides, in Chinese "Obstetrics and Gynecology", SOGC is adapted in antenatal fetal monitoring . In general, although the existing grading methods can define the specific conditions of the fetus, they have high sensitivity and low specificity in practical clinical applications [2−3,9,16]. Especially when the CTG case is less than 40 minutes, it is prone to false positives, which will lead to overdiagnosis of fetal distress and unnecessary cesarean section for pregnant women [17].

### 2.2 The machine-learning-based models for antenatal fetal monitoring

At present, many scholars at home and abroad mainly use machine learning methods to classify the CTG dataset of prenatal fetal monitoring research by Ayresde et al. in SisPorto 2.0 Portugal [18], through the use of uterine contractions (UC) and fetal heart rate (FHR). The properties of the data obtained by the signal are classified. In the literature [19] (2015) to evaluate the classification performance of eight different machine learning methods on prenatal CTG data, the study shows that the accuracy of the classifier is not much different. In the literature [20] Yang Zhang (2017) used hybrid PCA and AdaBoost to successfully classify CTG data and assess fetal status with an accuracy of 98.6%. In the literature [21] Vinayaka Nagendra (2017) used RF and

SVM to perform a three-class study of the fetal status, and the results showed that the accuracy was higher than 96%. In the literature [22] Zhao Z (2018) through the statistical test (ST), area under the curve (AUC) and principal component analysis (PCA) for feature selection, and then use three representative machine learning algorithm decision tree (DT) The support vector machine (SVM) and adaptive enhancement (AdaBoost) are classified into two categories. The experimental results show that the combination of AdaBoost and ST has strong classification ability, the accuracy is 92%, the sensitivity is 92%, and the specificity is 90%. In general, the current machine learning research methods for prenatal CTG feature interpretation are mainly classification algorithms, lifting algorithms and hybrid algorithms. The accuracy in the standard data set is not much different, and basically, good experimental results are obtained.

## 3 Material and methods

### 3.1 Dataset description

Based on the strict medical,the data comes from the SisPorto2.0,a program for automated analysis of cardiotocograms that closely follows the FIGO guidelines, analyses ante- and intrapartum tracings, performs no signal reduction, and has the possibility of simultaneously recording twins. SisPorto2.0 has been tested in over 6000 pregnancies. The system's FHR baseline was compared with an average of three experts' estimates, and the difference was under 8 bpm in all cases. A fair to the good agreement was found with experts' identification of accelerations, decelerations, contractions, and normal/reduced variability (proportions of agreement 0.64-0.89). This dataset contains 2126 Portuguese pregnant women CTG records from 29 to 42 gestational weeks. Each CTG record has 21 features, among the 2126 CTG records, record has 21 features,among the 2126 CTG records, 1655, 176 and 295 cases respectively belonged to the normal, suspicious or abnormal state.

From Fig.1, it was obvious that the proportion of the fetal status was not uniform, the sample in the normal state accounted for 78%, and the total of suspicious and abnormal samples only accounted for 22%. It could be seen that there was a serious classification imbalance in the CTG dataset.

### 3.2 SEMs for interpreting antenatal fetal monitoring

The relationship between observed variables and latent variables is expressed as a matrix equation.

$$x = \boldsymbol{\Lambda_x} \xi + \boldsymbol{\delta} \tag{1}$$

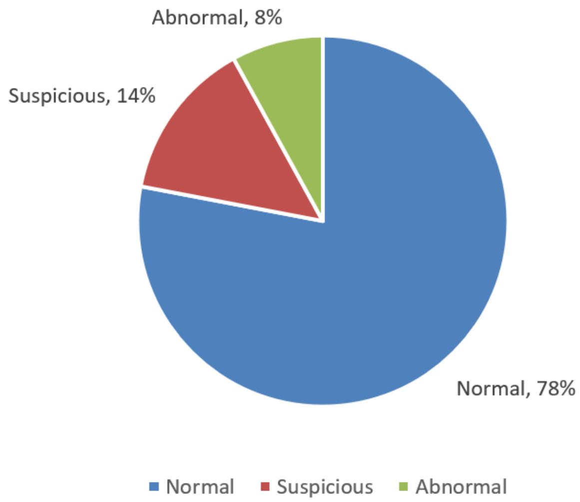


**Fig. 1.** CTG data distribution

$$y = \boldsymbol{\Lambda_y}\eta + \boldsymbol{\epsilon} \tag{2}$$

Where $x$ and $y$ are prenatal fetal monitoring features, $\xi$ and $\eta$ are latent variables(baseline category, variability category, acceleration category, deceleration category, uterine contraction), $\delta$ and $\epsilon$ are unique factor vectors. The simple form of the structural equation model is a multiple regression model with only one dependent variable. The formula is as follows, where $\eta$ is determined by y and $\epsilon$ .

$$\eta = \boldsymbol{\beta}\eta + \boldsymbol{\Gamma}\xi + \zeta \tag{3}$$

### 3.3 Intelligent antenatal fetal monitoring based on random forests

In the proposed weighted random forest model, the normal, suspicious and abnormal labels of CTG was defined as $y$ =0, 1 and 2, $n$ is total number of CTG cases. where $m, l$ and $n - m - l$ denoted the number of samples with labels 0,1 and 2; and $w_0, w_1$ and $w_2$ were the weights of categories 0, 1, and 2.

$$W_0 = \frac{n}{3m}, w_1 = \frac{n}{3l}, w_2 = \frac{n}{3(n-m-l)} \tag{4}$$

Based by the above, it can be seen that the penalty items of CTG categories were inversely proportional to the number of input samples. The larger the penalty items of a certain category, the higher the cost of misclassification would be considered. Hence, the WRF model is more sensitive to suspicious and abnormal categories

$$H(x) = argmax \sum_{k=1}^{n} \Pi(h_t(x) = y) \tag{5}$$

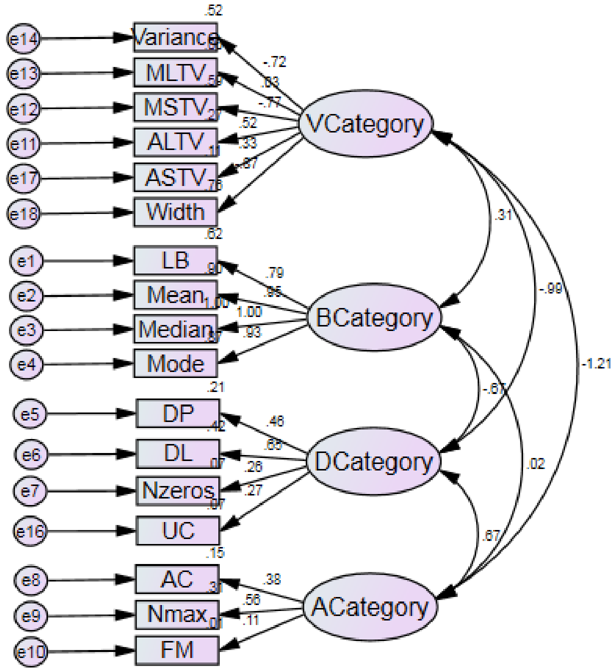


**Fig. 2.** CTG feature measurement model

After that, the WRF model randomly chose $K$ CTG features to obtain a subset $x_i$. Then, it would construct decision tree $h_i(Di)$, and made up random forest $h_1(D1), h_2(D2), h_3(D3)\ldots\ldots h_{i-1}(D_{i-1}), h_i(Di)$.Finally,the output prediction result was $H_1(x), H_2(x), H_3(x)\ldots\ldots H_{i-1}(x), H_i(x)$

## 4 Results and discussion

### 4.1 The experimental results of the SEMs models

#### *4.1.1 The result of measurement model assumptions*

With SisPorto 2.0: A program for automated analysis of Cardiotocograms, FIGO scoring, and clinical experience [23−25], combined with feature screening, classify 17 features into four categories as shown in Fig.2.

Table 1 reports the p-values of all factor loads are less than 0.001 (three stars), indicating that the interpretation of the measured variables (scale data) for the four latent variables (baseline, acceleration, deceleration, variant) is meaningful [32].

It can be seen from the p-value that the baseline and the acceleration class reach a significant level of 0.05. The estimated values of the baseline and variant classes are positive 22.463, indicating a correlation between the baseline and the variant of the hidden variable.

**Table. 1.** Estimated weighting factor

| Relationship | | | Estimate | C.R. | P |
|---|---|---|---|---|---|
| **LB** | ⟵ | BCategory | 0.789 | | |
| **Mean** | ⟵ | BCategory | 0.949 | 54.434 | *** |
| **Median** | ⟵ | BCategory | 0.999 | 58.593 | *** |
| **Mode** | ⟵ | BCategory | 0.934 | 53.124 | *** |
| **DP** | ⟵ | DCategory | 0.478 | | |
| **DL** | ⟵ | DCategory | 0.495 | 22.202 | *** |
| **AC** | ⟵ | ACategory | 0.328 | | |
| **FM** | ⟵ | ACategory | 0.177 | 6.929 | *** |
| **ALTV** | ⟵ | VCategory | 0.55 | | |
| **MSTV** | ⟵ | VCategory | -0.784 | -25.456 | *** |
| **ASTV** | ⟵ | VCategory | 0.376 | 15.154 | *** |
| **UC** | ⟵ | UC | 0.285 | | |
| **Nmax** | ⟵ | ACategory | 0.706 | 14.188 | *** |
| **Nzeros** | ⟵ | DCategory | 0.195 | 11.066 | *** |
| **Width** | ⟵ | VCategory | -0.855 | -26.586 | *** |
| **Variance** | ⟵ | VCategory | -0.704 | -23.95 | *** |

### *4.1.2 The result of prenatal CTG feature structure model*

The purpose of this model is to derive the causal relationship between the variable category (baseline category, variability category, acceleration category, deceleration category, uterine contraction) in the measurement model.

This structural model has no variables outside the model, and the variables that are interpreted are interpreted by variables within the model. Combined with the measurement model results, the model are obtained according to the model fit degree [26−28] which is shown in Fig.3.

The implicit factor causal relationship [33,34] correlation coefficient is shown in Table 2. From the p-value, the causal relationship of the three hidden factors has reached a significant level of 0.05. The model validates explain that the model hypothesis is established, the fetal heart rate contraction is verified to be the cause of fetal heart rate deceleration. The results also show that the variant class is the baseline class [35], which verifies the content of SisPorto 2.0: A Program for Automated Analysis of Cardiotocograms: the baseline value is adjusted according to the variation of the variant class to determine the value of the baseline of the heart rate.

In addition, the acceleration class and the deceleration class do not appear at the same time, so the Estimate value is quite high, but the acceleration does not appear to decelerate does not necessarily occur, so the S.E value is also large, and the knowledge of the fetal supervision is also consistent.

According to the fetal monitoring literature and books [23−25] , it is known

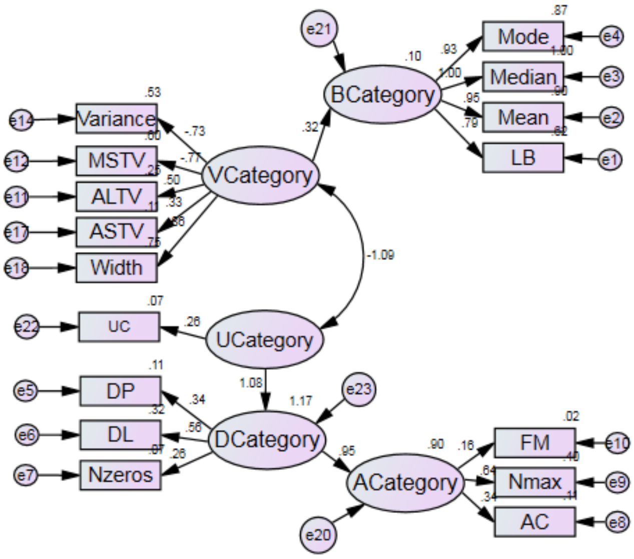


**Fig. 3.** Structural model of prenatal CTG features

that uterine contraction will block the blood flow between the uterus placenta, and the fetal oxygen supply will decrease, leading to a slowing of the fetal heart rate. When the rules are contracted, normal heart rate deceleration occurs, which is manifested as early deceleration. After the contraction slows down, the blood flow returns to normal, and the fetal heart rate quickly returns to normal. When the contractions are too strong or do not coordinate the uterine contractions, Cause fetal abnormalities such as late fetal deceleration or variability deceleration. So the model verifies that contractions are the cause of deceleration.

**Table. 2.** Implicit factor causality correlation estimate

| Relationship | | | Estimate | C.R. | P |
|---|---|---|---|---|---|
| **DCategory** | ⟵ | UC | 1.22 | 11.032 | *** |
| **ACategory** | ⟵ | DCategory | 0.891 | 13.188 | *** |
| **BCategory** | ⟵ | VCategory | 0.338 | 13.38 | *** |

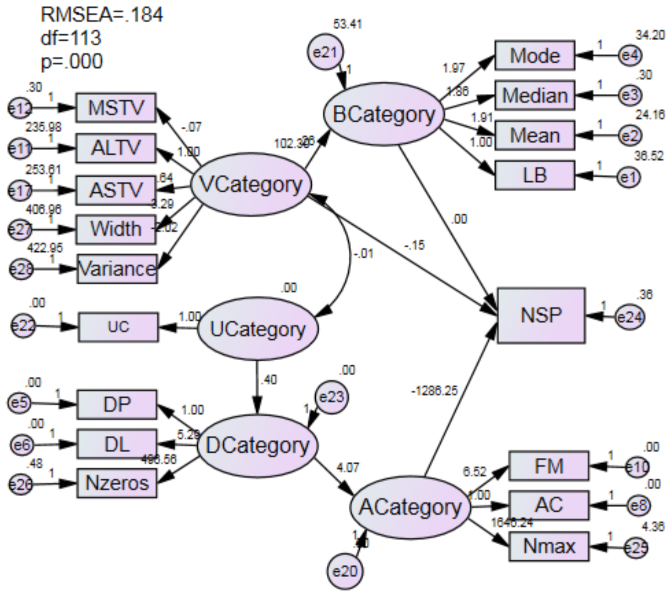


**Fig. 4.** Prenatal fetal state classification interpretative structural equation model

*4.1.3 The result of interpretative structural equation model for prenatal fetal status classification*

Prenatal fetal status NSP was added to the structural model of prenatal CTG features to explore the effects of four hidden factors on prenatal fetal status, providing interpretability for machine learning classification. According to the model fitting adjustment [29−31], the interpretable fetal state classification can be finally obtained as shown in Fig.4.

Table 3 reports the p-values of all factor loads are less than or equal to 0.001, indicating that the interpretation of the measured variables (scale data) by the four latent variables (baseline, acceleration, deceleration, variant) is meaningful. And the interpretation of the fetal state NSP by potential variables is also meaningful.

4.2 The experimental results of intelligent fetal monitoring model based on RF

With the outbreak of big data and artificial intelligence, the progress of the intelligent fetal evaluation methods is dramatically driven by computer science and engineering. Many scholars have researched on the intelligent fetal monitoring models based on machine learning. The sensitivity, specificity, F1

**Table. 3.** Correlation coefficient in the model

| Relationship | | | Estimate | C.R. | P |
|---|---|---|---|---|---|
| **DCategory** | ⟵ | UC | 1.22 | 11.032 | *** |
| **ACategory** | ⟵ | DCategory | 0.891 | 13.188 | *** |
| **BCategory** | ⟵ | VCategory | 0.338 | 13.38 | *** |
| **LB** | ⟵ | BCategory | 0.789 | | |
| **Mean** | ⟵ | BCategory | 0.949 | 54.434 | *** |
| **Median** | ⟵ | BCategory | 0.999 | 58.593 | *** |
| **Mode** | ⟵ | BCategory | 0.934 | 53.124 | *** |
| **DP** | ⟵ | DCategory | 0.478 | | |
| **DL** | ⟵ | DCategory | 0.495 | 22.202 | *** |
| **AC** | ⟵ | ACategory | 0.328 | | |
| **FM** | ⟵ | ACategory | 0.177 | 6.929 | *** |
| **ALTV** | ⟵ | VCategory | 0.55 | | |
| **MSTV** | ⟵ | VCategory | -0.784 | -25.456 | *** |
| **ASTV** | ⟵ | VCategory | 0.376 | 15.154 | *** |
| **UC** | ⟵ | UC | 0.285 | | |
| **Nmax** | ⟵ | ACategory | 0.706 | 14.188 | *** |
| **Nzeros** | ⟵ | DCategory | 0.195 | 11.066 | *** |
| **Width** | ⟵ | VCategory | -0.855 | -26.586 | *** |
| **Variance** | ⟵ | VCategory | -0.704 | -23.95 | *** |
| **NSP** | ⟵ | BCategory | -0.063 | -3.203 | 0.001 |
| **NSP** | ⟵ | VCategory | -2.557 | -8.358 | *** |
| **NSP** | ⟵ | ACategory | -2.687 | -7.871 | *** |

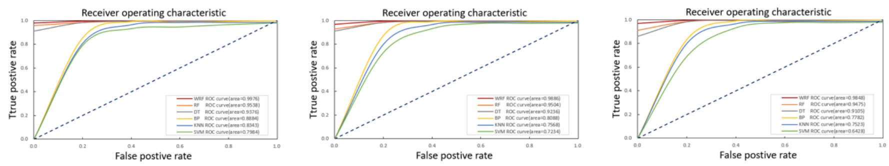


**Fig. 5.** CTG Multi-classification ROC curve

score and ROC curve area of the six machine learning models for prenatal fetal monitoring are shown in the table. Compared to other machine learning methods, the WRF model greatly improved sensitivity (0.99 and 0.98) and specificity (0.96 and 0.95) in the suspicious and anomalous categories. At the same time, according to the ROC curve area comparison chart (Fig.5), the WRF model is superior to other machine learning models in reducing the probability of misdiagnosis, and has the highest accuracy and the shortest running time. Therefore, the WRF model was chosen as the best representative of the machine learning model. Next, the SEM model was used to interpret the reason for the good performance of fetal status in fetal monitoring.

Compared with the existing antenatal CTG classification model (Table 4), the WRF model had best classification performance in imbalanced fetal monitoring data.

**Table. 4.** Comparison of overall accuracy and classification performance of existing CTG discriminant models

| Model | Feature | Normal | Suspicious | Abnormal | F1 | Accuracy |
|---|---|---|---|---|---|---|
| BP[36] | 22 | 97.84% | 45.14% | 97.24% | 80.07% | 91.31% |
| GRNN[37] | 21 | 95.70% | 73.92% | 84.88% | 84.83% | 91.86% |
| PNN[37] | 21 | 95.91% | 73.81% | 85.45% | 85.06% | 92.14% |
| MLPNN[37] | 21 | 95.00% | 68.43% | 80.50% | 81.31% | 90.36% |
| RF[38] | 21 | 96.40% | 79.60% | 91.20% | 89.07% | 93.60% |
| IAGA[39] | 6 | 96.83% | 79.15% | 89.41% | 88.46% | 93.89% |
| DT-AdaBoost[40] | 21 | 97.15% | 83.69% | 92.84% | 91.23% | 95.01% |
| DA[41] | 10 | 89.69% | 58.50% | 65.58% | 71.26% | 82.03% |
| LS-SVM-PSO-BDT[42] | 21 | 96.02% | 72.98% | 79.18% | 82.73% | 91.58% |
| DT[41] | 10 | 93.31% | 60.09% | 66.43% | 73.28% | 86.31% |
| RF | 10 | 96.53% | 77.36% | 89.16% | 87.68% | 93.43% |
| WRF | **10** | **99.75%** | **97.68%** | **95.24%** | **97.85%** | **99.71%** |

### 4.3 Interpreting prenatal fetal monitoring using SEMs models

The important features of the fetal state classification in the excellent representative WRF of the machine learning model are shown in Fig.6. The figure shows the importance of Median, Mean, Mode and LB in the Baseline category similar to the presence of ASTV, ALTV, MSTV in the Variation category. On the one hand, this can be explained by the larger estimate of the variation class Baseline category in the CTG feature structure model (22.463). On the other hand, the prenatal fetal state classification interpretative structural model concludes that the Baseline category has a smaller effect on fetal status discrimination, so the Variation category is more important than the Baseline category.

The correlation coefficient between Acceleration category and Deceleration category in the CTG feature structure model is large (0.677), but the predicted estimate is 0,which verify that there is a difference in the importance of the Acceleration category and the Deceleration category in the machine learning model.There is a huge difference between the Deceleration category and the fetal state in the prenatal fetal status classification interpretable structural model so that the relationship between Deceleration category and NSP cannot be smoothly converged in the model. More importantly, the model derives three hidden variables that contribute significantly to the determination of fetal status, verifying that machine learning yields feature categories that contribute significantly to fetal state discrimination.

## 5 Conclusion

The experimental results show that the variant category predicts the baseline category and that the contraction has a predictive effect on the decelera-

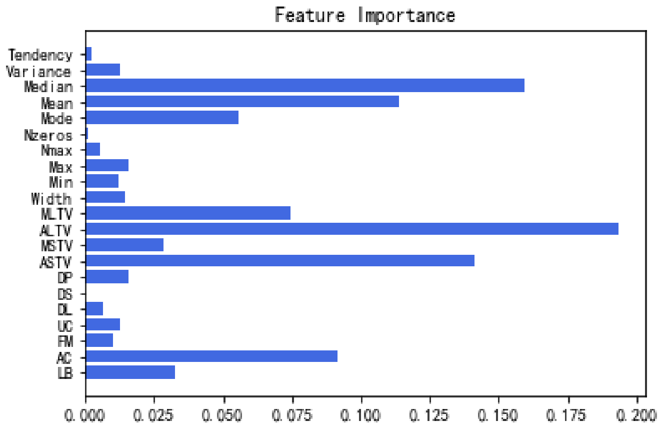


**Fig. 6.** CTG feature importance

tion category, validating the SisPorto 2.0 (CTG analysis program) and prenatal fetal monitoring clinical knowledge respectively. In addition, the variability category and the acceleration category have a greater impact on the identification of fetal status, while the baseline category has less impact, explaining the importance distribution of CTG features in weighted random forests.

**Acknowledgement :** This work is supported by the National Science Foundation of Guangdong Province under Grant No.2015A030310312, and the Medical Scientific Research Foundation of Guangdong Province under Grant No. B2014174.

## References

1. Haran S, Everett T R. Antenatal fetal wellbeing [J]. Obstetrics, Gynaecology and Reproductive Medicine,2017,27(2),44-49.
2. Grivell R M, Alfirevic Z, Gyte G M L, et al. Antenatal cardiotocography for fetal assessment[J]. The Cochrane Library,2015.
3. State Statistical Bureau. Statistical Yearbook of China[J]. Beijing: China Statistics Press,2017.
4. Kang-yue Li,Ning Ma.Artificial Intelligence and Its Application Status in Medical Field[J].Modern Medicine Health,2019,35(12):1823-1826.
5. Xia Fu. "Medical AI + rigorous medicine" is getting more and more attention [N]. China Insurance News, 2019-09-11 (005).
6. Zhang zhao. The Philosophical Basis and Interpretability of the Research Paradigm of Machine Learning [D]. Shanghai Academy of Social Sciences, 2018.
7. Ayresde C D, Bernardes J, Garrido A, et al. SisPorto 2.0: a program for automated analysis of cardiotocograms. [J]. Journal of Maternal-Fetal Medicine, 2000, 9(5):311-318.
8. Alfirevic Z, Devane D, Gyte G M. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour[M] The Cochrane Library. John Wiley & Sons,2006, Ltd:CD006066.
9. Cheng Zhihou, Song Shuliang. Fetal Electronic Monitoring [J].2001,81-92.

10. Lyons E R, Bylsma-Howell M, Shamsi S, et al. A scoring system for nonstressed antepartum fetal heart rate monitoring[J]. American Journal of Obstetrics & Gynecology,1979,133(3): 242-246.
11. Tian Caixuan. NST Score and Modified Fischer Scoring Method to Evaluate the Correlation between Fetal Test Results and Newborn Birth[J]. CJ,2009,16(22): 28-29.
12. Society of Obstetricians and Gynaecologists of Canada. SOGC clinical practice guidelines. Guidelines for vaginal birth after previous caesarean birth:155 (Replaces guideline Number 147), February,2015. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics,2005,89(3): 319.
13. 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 obstetricia et gynecologica Scandinavica,2017,96(2): 166-175.
14. Ayres-de-Campos D, Spong C Y, Chandraharan E. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography[J]. International Journal of Gynecology & Obstetrics,2015,131(1): 13-24.
15. Chinese Society of Perinatal Medicine. Expert consensus on the application of electronic fetal heart rate monitoring [J]. Chinese Journal of Perinatal Medicine,2015, 7: 486-490.
16. Xie Xing, Kong Beihua, Duan Tao. Gynecology and Obstetrics [M]. 9th Edition. Beijing. People's Health Publishing House,2018,54-56.
17. Zhang Huiping. Graphic analysis and management of 368 cases of abnormal non-stress test in fetal monitoring [J]. General nursing,2016,14 (36): 3837-3838.
18. Dua, D. and Karra Taniskidou, E. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science 2017.
19. Sahin, Hakan, and Abdulhamit Subasi. "Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques." Applied Soft Computing 2015,33,231-238.
20. Zhang, Yang, and Zhidong Zhao. "Fetal state assessment based on cardiotocography parameters using PCA and AdaBoost." 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE,2017.
21. Nagendra V, Gude H, Sampath D, et al. Evaluation of support vector machines and random forest classifiers in a real-time fetal monitoring system based on cardiotocography data[C]//2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB).2017, IEEE: 1-6.
22. Zhao Z, Zhang Y, Deng Y. A Comprehensive Feature Analysis of the Fetal Heart Rate Signal for the Intelligent Assessment of Fetal State[J]. Journal of clinical medicine, 2018, 7(8): 223.
23. Wei-hong Cheng, Zhi-hou Cheng . Fetal Electronic Monitoring (2nd Edition). People's Medical Publishing House, 2018
24. Hui Zhang, Qing-wen He, Shun-cang Zhang, et al. Prenatal ultrasound combined with fetal heart rate monitoring to predict the accuracy of intrauterine distress [J]. Journal of Guangxi Medical University, 2017 (2017 11): 1588-1590.
25. Hui-xia Yang, Xiao-tian Li, Zi-lian Wang, et al. Expert consensus on electronic fetal heart monitoring application[J]. Chinese Journal of Perinatal Medicine, 2015, 18(7): 486-490.
26. Browne, Michael W., and Robert Cudeck. "Alternative ways of assessing model fit." Sage focus editions 154 (1993): 136-136.
27. Zhonglin W, Kit-Tai H, Marsh H W. Structural equation model testing: Cutoff criteria for goodness of fit indices and chi-square test[J]. Acta Psychologica Sinica, 2004, 36(2): 186-194.
28. Bentler P M. On tests and indices for evaluating structural models[J]. Personality and Individual differences, 2007, 42(5): 825-829.
29. Browne, Michael W., and Robert Cudeck. "Alternative ways of assessing model fit." Sage focus editions 154 (1993): 136-136.
30. Zhonglin W, Kit-Tai H, Marsh H W. Structural equation model testing: Cutoff criteria for goodness of fit indices and chi-square test[J]. Acta Psychologica Sinica, 2004, 36(2): 186-194.
31. Bentler P M. On tests and indices for evaluating structural models[J]. Personality and Individual differences, 2007, 42(5): 825-829.

32. Byrne, Barbara M. Structural equation modeling with AMOS: Basic concepts, applications, and programming. Routledge, 2016.
33. Hernan, Miguel A., and James M. Robins. Causal inference. Boca Raton, FL: CRC, 2010
34. Peters J, Janzing D, Schölkopf B. Elements of causal inference: foundations and learning algorithms[M]. MIT press, 2017.
35. Koller D, Friedman N, Bach F. Probabilistic graphical models: principles and techniques[M]. MIT press, 2009.
36. Sundar C M, Chitradevi M C, Geetharamani G. Classification of Cardiotocogram Data using Neural Network based Machine Learning Technique [J]. International Journal of Computer Applications, 2013, 47(14):19-25
37. Yılmaz E. Fetal State Assessment from Cardiotocogram Data Using Artificial Neural Networks[J]. Journal of Medical & Biological Engineering, 2016, 36(6):820-832.
38. Arif M. Classification of cardiotocograms using random forest classifier and selection of important features from cardiotocogram signal[J]. 2015, 2(3):173-183.
39. Ravindran S, Jambek A B, Muthusamy H, 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,(2015-2-22), 2015, 2015.
40. Karabulut E M, Ibrikci T. Analysis of Cardiotocogram Data for Fetal Distress Determination by Decision Tree Based Adaptive Boosting Approach[J]. Journal of Computer & Communications, 2014, 02(9):32-37.
41. Huang M L, Hsu Y Y. Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network[J]. Journal of iomedical Science & Engineering, 2012, 05(9).
42. Y&X E, lmaz. Determination of Fetal State from Cardiotocogram Using LS-SVM with Particle Swarm Optimization and Binary Decision Tree[J]. Computational & Mathematical Methods in Medicine, 2013,(2):487179.