**Supplementary material**

**Supplementary Table 1.** Summary of all the publications for the last decade about the improve detection of perinatal outcome by the analysis of CTG data using ML methods.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author** | **Year** | **Dataset** | **Methodology** | **Best results** |
| **Accuracy** | **Sensitivity** | **Specificity** |
| Cömert Z. et al. [1] | 2019 | CTU-UHB1: 552 samples | Deep-CNN | 93.32% | 56.15% | 96.51% |
| Iraji M.S. [2] | 2019 | SisPorto2: 2126 samples | MLA-ANFIS; ANN; DSSAEs; deep-ANFIS | 99.50% | 99.72% | 97.50% |
| Sontakke S.S. et al. [3] | 2019 | SisPorto: 2126 samples | ANN; SVM, GLM; RF; KNN | 93.40% | - | - |
| Zhao Z. et al. [4] | 2019 | CTU-UHB: 552 samples (subset of 210) | CNN | 98.69% | 99.29% | 98.10% |
| Abbas R. et al. [5] | 2018 | CTU-UHB: 552 samples | GBM; GLMNET; KNN; SVM; RF; ANN | 95.40% | 87.50% | 100% |
| Fergus P. et al. [6] | 2018 | CTU-UHB: 552 samples | FLDA; RF; SVM | - | 87.00% | 90% |
| Zhao Z. et al. [7] | 2018 | CTU-UHB: 552 samples | DT; SVM; Adaboost | - | 92.00% | 90.00% |
| Abry P. et al. [8] | 2018 | CTU-UHB: 552 samples & LDB3: 3049 samples | SVM with spare learning | - | 62.00% | 89.00% |
| Petroziello A. et al. [9] | 2018 | CTU-UHB:552 samples & OxSys4: 35.429 samples | LSTM; CNN | - | 85.00% | 65.00% |
| Jianqian L. et al. [10] | 2018 | 4473 samples | CNN; SVM; MLP | 93.24% | 98.00% | 84.77% |
| Cömert Z. et al. [11] | 2018 | CTU-UHB: 552 samples | Least square-SVM | - | 63.45% | 65.88% |
| Cömert Z. et al. [12] | 2017 | SisPorto: 2126 samples | ANN; SVM; ELM; RBFN; RF | - | 99.73% | 97.94% |
| Georgoulas G. et al. [13] | 2017 | CTU-UHB: 552 samples | Least square-SVM | - | 68.48% | 77.68% |
| Barquero-Perez O et al. [14] | 2017 | 32 samples | KNN; SVM | - | 92.00% | 85.00% |
| Cömert Z. et al. [15] | 2016 | CTU-UHB: 552 samples (divided in stages I, II, III) | ANN  | 92.40% (I), 83.29% (II), 79.22% (III) | - | - |
| Cömert Z. et al. [16] | 2016 | CTU-UHB: 552 samples (subset of 100) | ANN | 87.00% | 88.70% | 85.10% |
| Spilka J. et al. [17] | 2016 | CTU-UHB: 552 samples & LDB: 3049 samples | SVM | - | - | - |
| Sahin H. et al. [18] | 2015 | SisPorto: 2126 samples | ANN; SVM; RBFN; KNN; CART; DT; RF | - | 94.18% | 99.74% |
| Chamidah N. et al. [19] | 2015 | SisPorto: 1831 samples | K-Means + SVM | 90.64% | - | - |
| Ibrahim O.I. et al. [20] | 2015 | 300 samples | RF; ruled-based and penalized RC | - | 99.74% | 86.00% |
| Peterek T. et al. [21] | 2014 | - | CART; RF | 94.69% | - | - |
| Ocak H. [22] | 2013 | SisPorto: 1831 samples | SVM + GA | 99.30% | - | - |
| Ocak H. et al. [23] | 2013 | SisPorto: 1831 samples | ANFIS | 96.60% | - | - |
| Georgieva A. et al. [24] | 2013 | 376 samples | ANN | - | 60.30% | 67.50% |
| Yilmaz E. et al. [25] | 2013 | SisPorto: 2126 samples | Least square-SVM | 91.62% | - | - |
| Czabanski R. et al. [26] | 2012 | 189 samples from 51 subjects | MLP; Lagrangian SVM + weighted fuzzy scoring system | 92.00% | - | - |
| Sundar C. et al. [27] | 2012 | SisPorto: 2126 samples | ANN | 60.00% | - | - |
| Warrick P.A. et al. [28] | 2012 | 264 patients | SVM | - | - | 92.50% |
| Mei-Ling H. et al. [29] | 2012 | SisPorto: 2126 samples | DA; DT; ANN | 97.78% | - | - |
| Krupa N. et al. [30] | 2011 | 90 samples from 15 subjects | SVM | 86.00% | 94.80% | 94.80% |
| Warrick P.A. et al. [31] | 2010 | 1315 samples | SVM | - | - | 92.5% |

1Open CTG database of 552 recordings from the University Hospital in Brno, Czech Republic [32].

2Open CTG database of 2126 recordings from the University of Porto, Portugal [33].

**Machine learning (ML) methods**: Adaboost = Adaptive boosting; ANFIS = Adaptive neuro fuzzy inference system; ANN = artificial neural network; CART = Classification and Regression Trees; CNN = Convolutional neural network; DA = discriminant analysis; DSSAE = Deep stacked sparse autoencoders; DT = decision trees; ELM = extreme learning machine; FLMA = Fishers Linear Discriminant Analysis; GA = Genetic algorithm; GBM = Gradient Boosting model; GLM= Generalized Linea Model; GLMNET = Generalized Linear Model Lasso and Elastic-Net Regularized; KNN = K-nearest neighbours; LR = logistic regression; MLA = multilayer architecture; MLP = multi-layered perceptron; RBFN = radial basis function network; RF = Random Forest; SVM = Support Vector Machine; TC: Tree classifiers.

**REFERENCES**

1 Cömert Z, Kocamaz AF. Fetal Hypoxia Detection Based on Deep Convolutional Neural Network with Transfer Learning Approach. Software Engineering and Algorithms in Intelligent System. Springer, Cham; 2019; pp 239–48.

2 Iraji MS. Prediction of fetal state from the cardiotocogram recordings using neural network models. Artif Intell Med. 2019 May;96:33–44.

3 Sontakke SA, Lohokare J, Dani R, Shivagaje P. Classification of Cardiotocography Signals Using Machine Learning. Proceedings of SAI Intelligent Systems Conference. Springer, Cham; 2019; pp 439–50.

4 Zhao Z, Zhang Y, Comert Z, Deng Y. Computer-Aided Diagnosis System of Fetal Hypoxia Incorporating Recurrence Plot With Convolutional Neural Network. Front Physiol. 2019 Mar;10:255.

5 Abbas R, Hussain AJ, Al-Jumeily D, Baker T, Khattak A. Classification of Foetal Distress and Hypoxia Using Machine Learning Approaches. International Conference on Intelligent Computing. Springer, Cham; 2018; pp 767–76.

6 Fergus P, Selvaraj M, Chalmers C. Machine learning ensemble modelling to classify caesarean section and vaginal delivery types using Cardiotocography traces. Comput Biol Med. 2018 Feb;93:7–16.

7 Zhao Z, Zhang Y, Deng Y, Zhao Z, Zhang Y, Deng Y. A Comprehensive Feature Analysis of the Fetal Heart Rate Signal for the Intelligent Assessment of Fetal State. J Clin Med. 2018 Aug;7(8):223.

8 Abry P, Spilka J, Leonarduzzi R, Chudáček V, Pustelnik N, Doret M. Sparse learning for Intrapartum fetal heart rate analysis. Biomed Phys Eng Express. 2018 Apr;4(3):034002.

9 Petrozziello A, Jordanov I, Aris Papageorghiou T, Christopher Redman WG, Georgieva A. Deep Learning for Continuous Electronic Fetal Monitoring in Labor. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2018; pp 5866–9.

10 Li J, Huang L, Shen Z, Zhang Y, Fang M, Li B, et al. Automatic Classification of Fetal Heart Rate Based on Convolutional Neural Network. IEEE Internet Things J. 2018;1–1.

11 Cömert Z, Kocamaz AF. Open-access software for analysis of fetal heart rate signals. Biomed Signal Process Control. 2018 Aug;45:98–108.

12 Cömert Z, Kocamaz AF. Comparison of Machine Learning Techniques for Fetal Heart Rate Classification. 3rd International Conference on Computational and Experimental Science and Engineering. 2017. DOI: 10.12693/APhysPolA.132.451

13 Georgoulas G, Karvelis P, Spilka J, Chudáček V, Stylios CD, Lhotská L. Investigating pH based evaluation of fetal heart rate (FHR) recordings. Health Technol (Berl). 2017 Nov;7(2–3):241–54.

14 Barquero-Pérez Ó, Santiago-Mozos R, Lillo-Castellano JM, García-Viruete B, Goya-Esteban R, Caamaño AJ, et al. Fetal Heart Rate Analysis for Automatic Detection of Perinatal Hypoxia Using Normalized Compression Distance and Machine Learning. Front Physiol. 2017 Feb;8:113.

15 Cömert Z, Fatih Kocamaz A. Evaluation of Fetal Distress Diagnosis during Delivery Stages based on Linear and Nonlinear Features of Fetal Heart Rate for Neural Network Community. 2016; [cited 2019 Jan 24].Available from: https://pdfs.semanticscholar.org/5e0b/a17a9da5ed9ec2543980dcddf5919941b5ef.pdf

16 Cömert Z, Fatih A. A Study Based on Gray Level Co-Occurrence Matrix and Neural Network Community for Determination of Hypoxic Fetuses. International Artificial Intelligence and Data Processing Symposium. 2016; [cited 2019 Jan 24].Available from: https://www.researchgate.net/profile/Zafer\_Coemert/publication/308684397\_A\_Study\_Based\_on\_Gray\_Level\_Co-Occurrence\_Matrix\_and\_Neural\_Network\_Community\_for\_Determination\_of\_Hypoxic\_Fetuses/links/582b0cf108aef19cb80696e1/A-Study-Based-on-Gray-Level-Co-Occur

17 Spilka J, Chudáček V, Huptych M, Leonarduzzi R, Abry P, Doret M. Intrapartum Fetal Heart Rate Classification: Cross-Database Evaluation. XIV Mediterranean Conference on Medical and Biological Engineering and Computing. Springer, Cham; 2016; pp 1199–204.

18 Sahin H, Subasi A. Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques. Appl Soft Comput. 2015 Aug;33:231–8.

19 Chamidah N, Wasito I. Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine. 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS). IEEE; 2015; pp 37–41.

20 Idowu IO, Fergus P, Hussain A, Dobbins C, Khalaf M, Eslava RVC, et al. Artificial intelligence for detecting preterm uterine activity in gynacology and obstertric care. Proc - 15th IEEE Int Conf Comput Inf Technol CIT 2015, 14th IEEE Int Conf Ubiquitous Comput Commun IUCC 2015, 13th IEEE Int Conf Dependable, Auton Se. 2015;215–20.

21 Peterek T, Gajdoš P, Dohnálek P, Krohová J. Human Fetus Health Classification on Cardiotocographic Data Using Random Forests. Advances in Intelligent Systems and Computing (AISC, volume 298). Springer, Cham; 2014; pp 189–98.

22 Ocak H. A Medical Decision Support System Based on Support Vector Machines and the Genetic Algorithm for the Evaluation of Fetal Well-Being. J Med Syst. 2013 Apr;37(2):9913.

23 Ocak H, Ertunc HM. Prediction of fetal state from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems. Neural Comput Appl. 2013 Nov;23(6):1583–9.

24 Georgieva A, Payne SJ, Moulden M, Redman CWG. Artificial neural networks applied to fetal monitoring in labour. Neural Comput Appl. 2013 Jan;22(1):85–93.

25 Yılmaz E, Kılıkçıer C. Determination of fetal state from cardiotocogram using LS-SVM with particle swarm optimization and binary decision tree. Comput Math Methods Med. 2013 Oct;2013:487179.

26 Czabanski R, Jezewski J, Matonia A, Jezewski M. Computerized analysis of fetal heart rate signals as the predictor of neonatal acidemia. Expert Syst Appl. 2012 Nov;39(15):11846–60.

27 Sundar C, Chitradevi M, Geetharamani G. Classification of Cardiotocogram Data using Neural Network based Machine Learning Technique. 2012; [cited 2019 Jan 9].Available from: https://pdfs.semanticscholar.org/5af7/3ff4d8297f901d89c8b7b9c42c26c11279fa.pdf

28 Warrick PA, Hamilton EF, Kearney RE, Precup D. A Machine Learning Approach to the Detection of Fetal Hypoxia during Labor and Delivery. Twenty-Second Innov Appl Artif Intell Conf. 2012;1865–70.

29 Huang M-L, Hsu Y-Y. Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network. J Biomed Sci Eng. 2012;5:526–33.

30 Krupa N, MA M, Zahedi E, Ahmed S, Hassan FM. Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. Biomed Eng Online. 2011 Jan;10(1):6.

31 Warrick PA, Hamilton EF, Precup D, Kearney RE. Classification of Normal and Hypoxic Fetuses From Systems Modeling of Intrapartum Cardiotocography. IEEE Trans Biomed Eng. 2010 Apr;57(4):771–9.

32 Chudáček V, Spilka J, Burša M, Janků P, Hruban L, Huptych M, et al. Open access intrapartum CTG database. 2014; [cited 2019 Jun 13].Available from: http://www.biomedcentral.com/1471-2393/14/16DATABASE

33 Ayres-de-Campos D, Sousa P, Costa A, Bernardes J. Omniview-SisPorto® 3.5 – a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. J Perinat Med. 2008 Jan;36(3):260–4.