

## On the Influence of Sleep Habits for Predicting Preterm Births: A Machine Learning Approach

**Authors :** C. Fernandez-Plaza, I. Abad, E. Diaz, I. Diaz

**Abstract :** Births occurring before the 37th week of gestation are considered preterm births. A threat of preterm is defined as the beginning of regular uterine contractions, dilation and cervical effacement between 23 and 36 gestation weeks. To author's best knowledge, the factors that determine the beginning of the birth are not completely defined yet. In particular, the incidence of sleep habits on preterm births is weekly studied. The aim of this study is to develop a model to predict the factors affecting premature delivery on pregnancy, based on the above potential risk factors, including those derived from sleep habits and light exposure at night (introduced as 12 variables obtained by a telephone survey using two questionnaires previously used by other authors). Thus, three groups of variables were included in the study (maternal, fetal and sleep habits). The study was approved by Research Ethics Committee of the Principado of Asturias (Spain). An observational, retrospective and descriptive study was performed with 481 births between January 1, 2015 and May 10, 2016 in the University Central Hospital of Asturias (Spain). A statistical analysis using SPSS was carried out to compare qualitative and quantitative variables between preterm and term delivery. Chi-square test qualitative variable and t-test for quantitative variables were applied. Statistically significant differences ( $p < 0.05$ ) between preterm vs. term births were found for primiparity, multi-parity, kind of conception, place of residence or premature rupture of membranes and interruption during nights. In addition to the statistical analysis, machine learning methods to look for a prediction model were tested. In particular, tree based models were applied as the trade-off between performance and interpretability is especially suitable for this study. C5.0, recursive partitioning, random forest and tree bag models were analysed using caret R-package. Cross validation with 10-folds and parameter tuning to optimize the methods were applied. In addition, different noise reduction methods were applied to the initial data using NoiseFiltersR package. The best performance was obtained by C5.0 method with Accuracy 0.91, Sensitivity 0.93, Specificity 0.89 and Precision 0.91. Some well known preterm birth factors were identified: Cervix Dilation, maternal BMI, Premature rupture of membranes or nuchal translucency analysis in the first trimester. The model also identifies other new factors related to sleep habits such as light through window, bedtime on working days, usage of electronic devices before sleeping from Mondays to Fridays or change of sleeping habits reflected in the number of hours, in the depth of sleep or in the lighting of the room. IF dilation  $\leq 2.95$  AND usage of electronic devices before sleeping from Mondays to Friday = YES and change of sleeping habits = YES, then preterm is one of the predicting rules obtained by C5.0. In this work a model for predicting preterm births is developed. It is based on machine learning together with noise reduction techniques. The method maximizing the performance is the one selected. This model shows the influence of variables related to sleep habits in preterm prediction.

**Keywords :** machine learning, noise reduction, preterm birth, sleep habit

**Conference Title :** ICAIMA 2018 : International Conference on Artificial Intelligence in Medical Applications

**Conference Location :** London, United Kingdom

**Conference Dates :** October 15-16, 2018