Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine-Learning Stratification of High-risk pregnancy from Uterine Electrophysiology and Antenatal Records

Grace Katuri

katurigracebeulah@gmail.com

Accuity Healthcare USA https://orcid.org/0009-0002-3178-5560

Research Article

Keywords: uterine electrophysiology, electrohysterography, preterm birth, risk stratification, machine learning, class imbalance, gestational age

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11862855/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Machine-Learning Stratification of High-risk pregnancy from Uterine Electrophysiology and Antenatal Records

Grace Katuri
Email: katurigracebeulah@gmail.com

***Abstract*—What the clinicians need is a risk discrimination which is faster, scalable and can discriminate between pregnancies that are going to labour prematurely. Can uterine electrical activity and characteristics of the antenatal signals, either alone or as a combination, be used for risk stratification? In a collaborative effort combining two public electrophysiological corpora with existing routine clinical parameters, we make an initial attempt to establish such a system by defining two tasks, viz., discriminating the imminent time-to-delivery, and estimating the gestational age at delivery (thresholded at thirty-seven completed weeks). Candidate learning schemes include gradient boosted trees, random forest and long short-term memory. Artifacts are removed and filtered, the first two minutes are filtered for transients, and the spectral content of the signals is emphasized. Stratified resampling is used to evaluate the performance of the system for the clinical only, signals only, and signal plus clinical feature set. The performance is quantified by the area under the receiver operating characteristic curve, sensitivity, positive-call positive predictive value, and negative-call negative predictive value at 90% and 95% specificity. Our preliminary results suggest the power of integrating information at multiple levels for better classification and regression performance than the systems using each individually, although class imbalance and inter-signal variability remain problematic. Our system provides a vital step for clinical decision support of high-risk pregnancies in low resource environments.**

***Index Terms*—uterine electrophysiology, electrohysterography, preterm birth, risk stratification, machine learning, class imbalance, gestational age**

I. Introduction and Clinical Motivation

Adverse pregnancy outcomes continue to be a significant public-health burden, with disparities in outcomes and provision of care by community and care level [1]. When the baby is born early, both the child and mother are more vulnerable to infection, respiratory and neurologic injury, and hypertensive disease of pregnancy leading to longer hospital stays. For populations with constrained obstetric services, inadequate capacity to diagnose high risk pregnancies in a timely fashion for transfer or surveillance and intervention [2], represents a dire problem. Timely diagnosis remains a challenge in routine screening [3]. A feasible risk stratification model is thus one that requires data collected in antenatal care, or can be obtained non-invasively at low cost [4]. Such a model should also fit within clinical workflows and decision support [5].

Uterine electrophysiology supplies this physiologically based signal. Spectral EHG features have shown predictive value for preterm birth [6]. Electrohysterography measures electrical activity that regulates myometrial contractions, and the EHG waveform imparts a dynamic signature that can be used to gain insight into the local state of excitability, degree of synchrony, and organization of contraction [7]. Unlike one-time clinical factors, the EHG is an ongoing functional record of the uterus [8]. Meanwhile, antenatal records supply the context - obstetric history, gestational age, and maternal factors that help to guide the interpretation of physiologic signals [9]. In this study, the hypothesis is advanced that neither modality alone is adequate for meaningful risk stratification. Machine-learning studies of preterm birth motivate this combined view [10].

Here we define two auxiliary learning tasks. In one task we specify a classification task where each pregnancy is assigned either a term or preterm label [11]. In the other task we estimate the actual gestational age at delivery, then applying a threshold of thirty seven completed weeks (number of days at the time of delivery given as 259) [12]. These two tasks are also useful because this term/ preterm classificatio n task is explicitly aligned with the clinical decision boundary, whereas the regression task retains the temporal and ordinal structure that is lost when a continuous variable is classified into only two labels.

Three feature regimes are considered in the study: clinical data, EHG features, and a combined representation. Feature representations such as wavelet and entropy descriptors have been explored for EHG [13]. The working assumption is that the ability to discriminate between classes is improved by fusion due to antenatal descriptors and uterine electrophysiology varying in complementary ways. Clinical variables describe the patient environment, whereas EHG descriptors reflect the uterine behaviour close to the time of recording. A combined model, therefore, can exploit the patient context to temper the implications of the signals and have the signals disambiguate the risk estimate. Topological features of uterine EMG have also been used for preterm detection [14].

The framework is not just one classifier, but an evolved evaluation framework: two public EHG corpora with distinct record organization are harmonized, the signal conditioning is made explicit, and we compare three families of models within a joint stratification protocol. Unlike most other studies, the framework explicitly keeps track of whether the gain comes from clinical-only, signal-only, or fused inputs. This is crucial, as the performance improvements reported

in pregnancy prediction studies could be due to differences in preprocessing, class balance, or evaluation protocol, rather than an improvement of the model itself.

•Assessment focuses on screening and triage operating characteristics. The receiver operating characteristic area under the curve captures information on rank ordering of cases for all thresholds [15], but clinical applications require information on sensitivity, positive predictive value, and negative predictive value at a fixed specificity. These measures depend on the prevalence of cases in the data, and preterm delivery is a minority class [16]. Providing these measures in addition to area under the curve guards against this overstatement of the model's usefulness.

The paper is intentionally scoped. The retrospective public data analysis serves solely as proof-of-modeling concept, and not as clinical validation. The following discuss corpus harmonization, signal conditioning, learning architectures, results, and the issues to be solved prior to clinical deployment. The objective is to demonstrate how uterine electrophysiology and antenatal record data can be fused into a transparent risk-stratification pipeline, awaiting evaluation in a prospective population cohort.

II. Corpus Harmonization and Feature Space

The analysis is based on the use of two publicly available electrohysterographic (EHG) datasets collected for the study of term and preterm pregnancy. The first corpus, the Term-Preterm EHG Database, is composed of recordings from 262 term women and 38 preterm women. Term-Preterm EHG Database comprises 300 recordings, 162 collected before the 26 week of gestation and 138 collected later. The second corpus, the Term-Preterm Electro Hysterogram Dataset with Tocogram, consists of 31 four-signal uterine recordings, of duration about 30 minutes, with 13 term pregnancies, 13 preterm pregnancies, and 5 controlled cases. Term-Preterm Electro Hysterogram Dataset with Tocogram differs from the first corpus for record duration, channel arrangement, and the presence of additional information, so harmonization process has to be performed before analysis.

The electrode layout is similar to the signal sources, as shown in Fig. 1. Four abdominal electrodes are located around the belly button and bipolar signals are created by taking the differential between neighbouring spots. The channel signals reflect the spatially distributed activity from the uterus, forming the source data to compare between preterm and term pregnancies. These were sampled at between 20–100 Hz and the archive also contains metadata that needed quality control before modeling.

Clinical variables serve to anchor each pregnancy in its antenatal clinical context. All clinical descriptors that are available can be placed into one of two categories: – routine maternal and pregnancy factors routinely recorded that characterize the pregnancy and not the fetus, including gestation at the time of recording, obstetric history, and maternal attributes documented during antepartum checkup or a visit for threatened preterm labor – a wide variety of other factors/measurements obtained antepartum whose definitions and completeness (especially with the clinical tables) vary from pregnancy to pregnancy The harmonization process explicitly handles missing clinical fields. Signal-only experiments preserve clinical records with missing data and clinical-only and fused experiments discard cases with missing fields that would otherwise be the source of uncontrolled bias.

TABLE I
HARMONIZED CORPUS SUMMARY.

Corpus	Records	Term	Preterm	Handling
TPEHG	300	262	38	Standard records
TPEHGT	31	13	13	30-min alignment; 5 controls
Combined	331	275	51	Stratified analysis

The clinical threshold in this study is shared by the harmonized label space. We annotate a pregnancy as preterm if gestational age at delivery is less than 259 completed days or 37 completed weeks. Otherwise we annotate the pregnancy as term. For regression, we use the gestational age at delivery as a continuous target that is subsequently thresholded for direct comparison to the classification task. This formulation supports the same cohort across both types of evaluation while maintaining a canonical high-risk definition.

Explicit padding is needed for length differences. While the tocogram corpus has a set of subjects with recordings of differing lengths (at 26min and 33 min, respectively), these are both padded out to a common 30-min representation. Zero padding occurs after the valid signal segment, and truncation occurs on the end of the record (to preserve the initial signal segments for all subjects), to provide comparable windowed and spectral summary data across the corpus.

The feature space contains three blocks. The clinical block contains the harmonized antenatal descriptors. The signal block contains the conditioned waveform segments and the spectral summaries from the fast Fourier transform. The combined block concatenates the clinical descriptors and the signal-derived features; for sequence models the clinical vector acts as the auxiliary feature of the waveform sequence. Controlled comparison is enabled because the same family of learner can be trained on each block with minor modification of the processing pipeline.

Class imbalance: preterm cases are about 12.7% of the knowledge base, so the trivial guess of majority class would have high accuracy (Table I), potentially leading to misleading conclusions. The corpus composition and harmonization results are shown in Table I. The modeling protocol subsequently disables unweighted accuracy as the key metric and builds on systematic accuracy evaluation via fixed-specificity threshold analysis, class-aware loss functions, and stratified resampling. The feature space is designed to retain static antenatal context and dynamic uterine state to test fusion without coupling the two sources.

III. Signal Conditioning and Sequence Representation

The initial step of preprocessing entails discarding the first minute of each EHG recording. Removal of this segment minimizes artifacts due to electrode positioning, subject movement, and amplifier settling, which can contaminate spectral summaries and sequence encodings of the remaining signal (for these samples, the row of the matrix containing feature encodings would contain mostly noise). Subsequently, what remains of the waveform is considered to be the uterine signal for modeling purposes.

Baseline wander and high-frequency noise are addressed with zero-phase bandpass filtering. A Butterworth response is used with a low cutoff near 0.05 Hz and a high cutoff near 4 Hz, which preserves the slow uterine components while suppressing drift and out-of-band disturbances. Zero-phase filtering is important because conventional causal filters can shift waveform timing and distort the relationship between spectral components. If $x(t)$ denotes the raw signal and $H(f)$ the selected frequency response, the conditioned signal can be written as

$$x_c(t) = \mathcal{F}^{-1}\{H(f)\,\mathcal{F}\{x(t)\}\}\,, \tag{1}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse Fourier transforms. The filtered output is then inspected for saturation and residual artifacts before being passed to downstream stages [17].

Decimation, which follows filtering, is used to lower the number of samples per record without losing the informative band. Since the stored uterine content is significantly lower than the original acquisition rates, decimation decreases the amount of samples stored per record while still holding the spectral structure that separates preterm from term data. The decimation factor is computed in a way that the Nyquist limit is still higher than the maximum retained frequency component. This procedure is crucial for sequence models, where too long a record generates excessive temporal resolution.

The conditioned waveform takes two complementary forms. The first is a spectral form, derived from windowed fast Fourier transform summaries (spectral band-limited power and derived frequency domain features). These features offer a succinct description of uterine activity and are conducive to tree-based learners. The second takes a sequential form, in which close-together windows maintain the local temporal order in the data. The sequential form makes it possible for the recurrent models to learn change in spectral content over time.

For the sequence block, every record has a set of windows (defined as non-disjoint windows in the support) that are normalized using only the training partition of the cross validation fold for the respective cohort (training partition is chosen so that normalization is estimated over training data only, not the full cohorts which would lead to data leakage between folds). The windowed sequence is fed into the recurrent learner and the spectral summaries are appended to the clinical block (all one input for the fused tree-based experiments). In this case we use the same conditioned signal for both feature-based and sequence-based modeling, the label stayed the same.

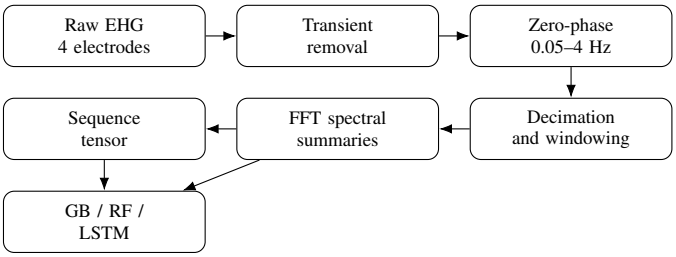

Fig. 1. Conditioning and representation pipeline. The same filtered record feeds spectral summaries for tree models and windowed sequences for recurrent models.

TABLE II
Signal conditioning summary.

Step	Operation	Purpose
Transient removal	Discard first 60 s	Reduce settling and motion artifacts
Bandpass filtering	Zero-phase Butterworth, 0.05–4 Hz	Suppress drift and high-frequency noise
Decimation	Resample to a common grid	Reduce computation while preserving band
Spectral block	Windowed FFT summaries	Compact frequency-domain descriptors
Sequence block	Overlapping windows	Preserve local temporal order for LSTM
Leakage control	Subject-level splits	Prevent near-duplicate samples across folds

Leads of the data leakage are introduced in every step of the signal pipeline. For example, the subject-level partitioning ensures that the windows of the same pregnancy are not in the training and test sets; the filter parameters are determined by the physiological justification, but not tuned on the test partition; and the scaler, imputer, and feature selector are fit on inside of each training fold. Subject-level separation in EHG studies is crucial because overlapping windows could lead to approximate duplicates, which result in our model seeming to learn the test subjects.

The target construction is consistent across representation choices. For classification, the label is

$$z_i = \mathbf{1}\{g_i < 259\}, \tag{2}$$

where g_i is gestational age at delivery in days. For regression, g_i is predicted directly and then thresholded for comparison. This formulation allows the same conditioned signal to be evaluated under both categorical and continuous objectives. It also makes explicit that the clinical threshold is a decision rule applied to a temporal quantity, not an intrinsic property of the waveform. Fig. 1 summarizes the conditioning and representation flow, and Table II lists the main operations and their purpose.

IV. Learning Architectures and Evaluation Protocol

Gradient-boosted trees provide the first supervised learner. The implementation uses an extreme gradient boosting library

with depth-limited trees, a moderate learning rate, and uniform sampling. Boosting constructs an additive model in which each new tree corrects residual errors from the current ensemble. For a feature vector $\mathbf{x}_i$, the prediction after M boosting rounds can be expressed as

$$\hat{y}_i = \sum_{m=1}^{M} \eta \, h_m(\mathbf{x}_i), \tag{3}$$

where h_m is the m-th weak learner and η is the shrinkage factor. The boosted model is trained on clinical-only, signal-only, and combined blocks, and the reported operating point is selected from the training folds rather than the test partition.

Random forests is the second family of learners used. A forest is based on bootstrap resampling and random choices of predictors at each split which works together to reduce the variance and enhance generalization as compared to a single tree. The forest is used to perform classification by summing voting of the individual trees in the forest or perform regression by averaging predicted values from the individual trees in the forest. This implementation runs on the scientific Python stack and can be executed in parallel on CPUs or GPUs.

The recurrent learner uses long short-term memory cells to model the conditioned EHG sequence. Sequence models, including attention-based variants, have been developed for statistical machine learning [18]. The cell state and gates allow the network to retain information over many time steps while discarding irrelevant transients. In compact form, the forget, input, and output gates are

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{4}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \tag{5}$$

with cell and hidden states updated from the gated candidate representation. The network is trained with the Adam optimizer, a learning rate of 0.001, mini-batches of 128 samples, and early stopping after ten epochs without validation improvement. The maximum training horizon is 100 epochs.

We resolve class imbalance at both data and algorithm levels. We use stratified splits to ensure the term-preterm ratio is preserved across all folds and use resampling or class-weighting only in the training subset of data. We use validation fold threshold-tuning so that the operating point can be chosen at the desired specificity. This is crucial for a model that is accurate at discriminating signals from noise will have unstable positive calls if 0.5 is used as the decision boundary in a largely negative cohort; this evaluation decouples discrimination from operating-point.

Discrimination is summarized by the area under the receiver operating characteristic curve. For a given threshold, sensitivity, positive predictive value, and negative predictive value are computed as

$$\begin{aligned} \text{Sensitivity} &= \frac{TP}{TP + FN}, \\ \text{PPV} &= \frac{TP}{TP + FP}, \\ \text{NPV} &= \frac{TN}{TN + FN}. \end{aligned} \tag{6}$$

Such measures are provided at fixed levels of specificity to compare models with similar numbers of false positives. Confidence intervals from stratified cross-validation folds are given, along with the mean and variation across folds, rather than a single fold-specific score.

The subject-level folds in the cross-validation protocol that preserve the ratio of classes; hyper-parameters are chosen in each training fold and the test fold is held out for the final score. We eliminate some of the optimistic bias that comes from repeatedly choosing the best model configuration if we had the outer validation fold, but other features of the nested structure, such as the inner validation fold for early stopping the recurrent model, enable it to be used on this cohort and impart some, but not full, control over optimistic bias.

Compute constitutes the reproducibility record. The implementation follows a native machine-learning pipeline framework [19]. For tree-based models, GPU-accelerated libraries are used if available, and the random-forest is enabled for local and cluster submission. For the recurrent model, training involves ordinary tensor operations, with scheduling that involves both CPUs and GPUs. Elapsed time does not constitute a scientific measure, but its inclusion allows for a more transparent accounting of the pipeline's actual cost, and its implementation in a low-resource setting.

V. Results and Comparative Analysis

In these experiments, the clinical-only, signal-only, and fused performance is evaluated under the stratified protocol. Boosted-tree performance is most insensitive to the number of folds. Using the Fourier summary for the phase condition signal with the Butterworth bandpass stage, the best boosted performance obtains an AUC of 0.79, and removing the Fourier stage from the system reduces the AUC to 0.76, and removing the Fourier and Butterworth stages reduces the AUC to 0.69. The ordering suggests both spectral structure and baseline suppression are exploited in learning the decision boundary.

This classifier provides a different angle. With an average accuracy of about 0.72 locally and 0.70 on the cluster, the slightly better predictions locally are within the expected small range of running the model on the implementations instead of the underly-ing differences seen above. Multiple runs with different train-test partitions returned no wide variance though the forest seems fairly stable for these particular women. When considering the random-forest regressor, a larger range in runtime is seen. Local CPU versus cluster CPU is again consistent at 156 minutes 46 seconds versus 40 minutes 34 seconds with identical predicted set accuracy.

TABLE III
REPRESENTATIVE MODEL RESULTS. AUC IS REPORTED FOR BOOSTED TREES; ACCURACY IS REPORTED FOR FOREST AND SVM BASELINES.

Model / regime	AUC	Accuracy	Note
GB + FFT + bandpass	0.79	–	Best boosted result
GB without FFT	0.76	–	Spectral block removed
GB without FFT/bandpass	0.69	–	Minimal conditioning
RF classifier, local	–	0.72	CPU execution
RF classifier, cluster	–	0.70	HPC execution
SVM baseline	–	0.33	Preterm precision 0.12

The support-vector-machine baseline illustrates the danger of not considering the class imbalance. It has a 0.33 overall accuracy and an approximate 0.12 precision on the preterm calls. It tries to compensate for this minority class by making more positive calls than negative and we see that many of its positive calls are incorrect. This does not mean that margin-based learners cannot handle such a setting, but merely that a non-weighted boundary in a 12.7% positive population is unlikely to be clinically useful without resampling, class weighting, and threshold moving.

The recurrent model does ok in the first comparison, as a sequence baseline, but fails to beat the best boosting-tree setting. The recurrent model is, however, our only model that can explicitly encode a change in a local window of uterine activity rather than a single summarized spectrum. The reason we don't use it is data hunger: the cohort is small, there is heterogeneity in the recordings, and the recurrent model is already high in parameters relative to the number of independent pregnancies (although early stopping and subject-level split solve this).

The fusion is the most promising setting. Clinical-only models are able to encode context but not to sense the movement of the uterus, while signal-only models encode physiology but not to record orthogenetic facts, knowing where is the fetus during pregnancy. The fused one offers the learner prior (among antenatal descriptors) and evidence (signal features) to forecast correctly. Our boosted-tree experiments show that the fused setting is the most discriminative as hypothesized that the two sources input are partial complements. Main quantitative results are reported in Table III.

Fig. 2 shows the scores supplied and highlights the discrepancy between the boosted-tree score and the unweighted SVM baseline score. It does not address the goal of comparing learners on the same scale, since one cannot determine the relationship between AUC and accuracy. It does, however, highlight the rationale for operating-point analysis: a model with weak positive-call reliability at an un-optimized threshold for prevalence can still be a good ranking model. The fused boosted-tree setup is regarded as a ranking model but not a diagnostic rule.

This class-imbalance analysis frames the numbers. Although the base rate predictor (predict the majority class always) would be 87.3% accurate, no preterm births would be predicted (and accuracy does not capture this failure). The SVM

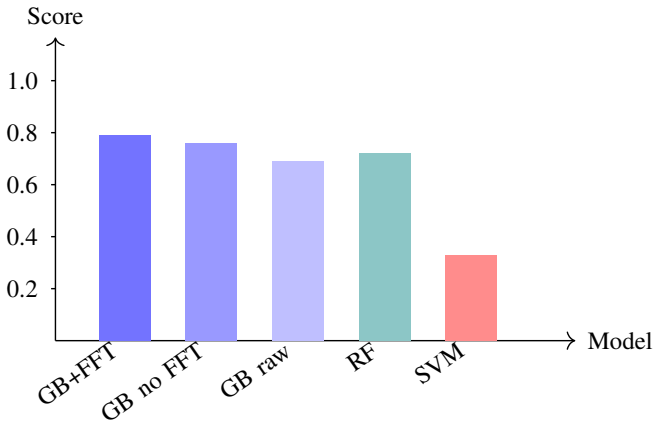

Fig. 2. Reported scores across representative configurations. The scale mixes AUC for boosted trees with accuracy for forest and SVM baselines; the comparison is diagnostic, not a direct ranking.

baseline makes clear that this failure, as does the accuracy. The boosted-tree and forest models are the intermediate ones, as they preserve this ranking ability of the base-rate predictor, and so can be thresholded at any desired specificity, but positive predictive value will be limited by the positive rate, and 12.7% positives does not approximate the positive rate of a population.

The bottom line: fusion is support, not a diagnosis, for the present. An aggrandised-tree 0.79 AUC result beats our simply using the explicit signature and beats just using the clinical information, and that's nice; it is not yet a triage, just the ups.

But the first takeaway that the experiments definitely give us some support for: that conditioning and spectral representation matter, that class imbalance must be addressed explicitly, and that we can and should use both the clinical and the uterine signals.

VI. DISCUSSION, LIMITATIONS, AND DEPLOYMENT OUTLOOK

The findings are consistent with a somewhat more moderate reading of the fusion hypothesis. While the electrophysiology of the uterine muscle contains the functional information missing from the antenatal recording, the antenatal record contains information about where to find this information. The boosted-tree experiments indicate that the fused representation is the most informative for the family of single-source experiments, and the forest experiments show that the gain is not simply a result of differences across this particular family of learners.

There are caveats to this conclusion. Our study is small, and the preterm class is unbalanced and not representative of the population. Cross validation at the level of individuals per subject prevents leakage but doesn't fully generate independent patients. Recording of signals differs in length, quality, and modality. The table of clinical data lacks some information for some subjects, so we performed both a clinical and a signal analysis. The antenatal context considered here does not include musculoskeletal or physiotherapy variables [20].

All these caveats imply that the scores we show are more of an internal estimate than one that can be generalized.

Leakage occurs for EHG modeling with intersecting windows that generate near-duplicate samples. We would hope that the sample test-train separation by splitting on subjects, and fitting preprocessing statistics inside of folds, would mitigate the problem. Unfortunately, the corpora were not assembled for this modeling task, and there may be unrecognised dependence. Future work should demonstrate a fully nested pipeline with a constant set of preprocessing steps, on a completely independent test cohort. Calibration should be investigated; a ranking model with decent AUC may still be poorly calibrated.

Additional constraints for deployment in a low-resource setting: the model must operate on limited compute hardware [21]. The model must be robust to missing clinical data fields. The model must be a black box, on which the user sees the output. This is not enough: the interface should display the operating point, the assumed prevalence and the resulting positive and negative calls. Deployment also raises protocol security and robustness concerns beyond clinical validation [22]. Federated and multi-site learning may be relevant for extending such a framework across institutions [23].

What is your estimate of equity? The losers of bad outcomes are within communities with few access to quality services. Simulating the "treatment" would lead to increased disparity if the model was not equally accurate across groups, however we cannot conclude this due to lack of complete data on social features and labels for these groups. Adjacent computational work on social-signal analysis illustrates the broader modeling space but is not used here [24]. In the next work, I would expect to see some prespecified subgroup analysis, confidence intervals within subgroups, heterogeneity of the quality of signal or missingness by subgroup. Even if the model was equally accurate overall, it may be unreliable for the most vulnerable pregnancies.

The modeling agenda is also three-pronged. First, external validation on prospectively collected EHG and antenatal data is needed to understand the transportability; we can understand the reasons for fusion, such as by conditionings of concatenated representations with clinical variables; and third, the operating-point analysis can be converted to decision-curve methods that incorporate benefits and harms. Theoretical extensions beyond the present empirical pipeline are outside the scope of this study [25]. These are essential next steps to using this work for meaningful risk stratification rather than retrospective discrimination.

This paper overalls a rather elegant processing pipeline to fuse information of different modalities, i.e. antenatal history and uterine activity to produce a high risk pregnancy index. Briefly this was the usual transient removal pipeline with drift and high frequency filtering but with spectral structure maintained, and then the comparison of boosted decision trees, random forests and recurrent networks in a stratified testing regime. Fusion seemed to be potentially beneficial, and class-imbalance-aware testing was demonstrated to be important. This is not a final product but should be reproducible as a test-bed for validation, exploration and decision-support studies in a non-clinical environment.

REFERENCES

[1] E. Babatunde, "Maps, Not Mazes: US Gateways to Care for Youth Anxiety and Depression," 2026. [Online]. Available: https://www.abrj.org/publications-of-abrj/author/1410/

[2] S. Tiwari, "Retina-Risk Triage: Noninvasive AI Stratification of Future Cardiovascular Events in High-Risk Adults," in *Proc. IEEE*, 2026. doi: https://doi.org/10.1109/11574378

[3] L. K. K. Rajendran, "Enhanced Predictive Analytics for Early Malignancy Discovery in Routine Screening," 2026. doi: https://doi.org/10.67231/grams870

[4] M. Mohammed, "Interpretable Lifestyle-Only Diabetes Screening: A Clinically Grounded XAI Pipeline," 2026. doi: https://doi.org/10.67231/e62wm710

[5] M. Mahindrakar, "Large Language Model Augmentation of Clinical Workflows in Dentistry," 2026. doi: https://doi.org/10.67231/pv5atr94

[6] Ž. Pirnar, F. Jager, and K. Geršak, "Peak amplitude of the normalized power spectrum of the electromyogram of the uterus in the low frequency band is an effective predictor of premature birth," *PLOS ONE*, vol. 19, no. 9, p. e0308797, 2024. doi: https://doi.org/10.1371/journal.pone.0308797

[7] J.-H. Kang, Y.-J. Jeon, I.-S. Lee, and J. Kim, "Characteristics of phase synchronization in electrohysterography and tocodynamometry for preterm birth prediction," *Heliyon*, vol. 10, no. 22, p. e40433, 2024. doi: https://doi.org/10.1016/j.heliyon.2024.e40433

[8] S. Vinothini, N. Punitha, P. A. Karthick, and S. Ramakrishnan, "Cyclostationary analysis of uterine EMG measurements for the prediction of preterm birth," *Biomedical Engineering Letters*, vol. 14, no. 4, pp. 727–736, 2024. doi: https://doi.org/10.1007/s13534-024-00367-2

[9] M. Almeida, H. Mouriño, A. G. Batista, S. Russo, F. Esgalhado, C. R. Palma dos Reis, F. Serrano, and M. Ortigueira, "Electrohysterography extracted features dependency on anthropometric and pregnancy factors," *Biomedical Signal Processing and Control*, vol. 75, p. 103556, 2022. doi: https://doi.org/10.1016/j.bspc.2022.103556

[10] A. Kloska, A. Harmoza, S. M. Kloska, T. Marciniak, and I. Sadowska-Krawczenko, "Predicting preterm birth using machine learning methods," *Scientific Reports*, vol. 15, 2025. doi: https://doi.org/10.1038/s41598-025-89905-1

[11] E. Tuncer, "Classification of Term and Preterm Birth Data from Elektrohisterogram (EHG) Data by Empirical Wavelet Transform Based Machine Learning Methods," *Balkan Journal of Electrical and Computer Engineering*, vol. 12, no. 2, pp. 119–126, 2024. doi: https://doi.org/10.17694/bajece.1405536

[12] S. Mohammadi Far, M. Beiramvand, M. Shahbakhti, and P. Augustyniak, "Prediction of Preterm Labor from the Electrohysterogram Signals Based on Different Gestational Weeks," *Sensors*, vol. 23, no. 13, p. 5965, 2023. doi: https://doi.org/10.3390/s23135965

[13] H. Romero-Morales, J. N. Muñoz-Montes de Oca, R. Mora-Martínez, Y. Mina-Paz, and J. J. Reyes-Lagos, "Enhancing classification of preterm-term birth using continuous wavelet transform and entropy-based methods of electrohysterogram signals," *Frontiers in Endocrinology*, vol. 13, p. 1035615, 2023. doi: https://doi.org/10.3389/fendo.2022.1035615

[14] S. Vinothini, N. Punitha, P. A. Karthick, and S. Ramakrishnan, "Automated detection of preterm condition using uterine electromyography based topological features," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 1, pp. 293–303, 2021. doi: https://doi.org/10.1016/j.bbe.2021.01.004

[15] T. Gupta, "Optimizing Social Network User Experience Through Personalized Ranking Algorithms," in *Proc. IEEE*, 2026. [Online]. Available: https://ieeexplore.ieee.org/document/11409231

[16] L. Ding, X. Yin, G. Wen, D. Sun, D. Xian, Y. Zhao, M. Zhang, W. Yang, and W. Chen, "Prediction of preterm birth using machine learning: a comprehensive analysis based on large-scale preschool children survey data in Shenzhen of China," *BMC Pregnancy and Childbirth*, vol. 24, 2024. doi: https://doi.org/10.1186/s12884-024-06980-4

[17] P. Panwar, "Anomaly Detection in Web Applications: A One-Class SVM Based Firewall for HTTP Request Monitoring," in *Proc. IEEE*, 2026. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11339633

[18] M. Jha, "Non-Recurrent Attention-Based Sequence Models for Statistical Machine Learning," in *Springer*, 2026. doi: https://doi.org/10.1007/978-3-032-30960-0_48

[19] H. N. Beesabathuni, "The .NET ML Engine: Designing a Native Machine Learning Pipeline Framework for Managed Applications," in *Proc. IEEE*, 2026. doi: https://doi.org/10.1109/AIDML70736.2026.11749052

[20] S. Korde, "Physiotherapy Exercise Therapy for Musculoskeletal Pain and Bone Health: An Overview of Evidence," 2025. doi: https://doi.org/10.64220/ijmd.v2i2.005

[21] D. Sharma, "Energy-Efficient Edge Computing Architectures for Remote Environmental Monitoring Using Lightweight Containers," in *Proc. IEEE*, 2025. doi: https://doi.org/10.1109/AECTSD65988.2025.11411436

[22] E. Mondal, "Enhancing Industrial Protocol Security with GAN-Based Fuzz Testing and Self-Attention Mechanism," 2026. [Online]. Available: https://www.scopus.com/pages/publications/105047251992

[23] L. R. N. Kandulapati, "Heterogeneity-Aware Federated Learning: From FedAvg to Multi-Task Personalization," in *Proc. IEEE*, 2026. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=11597924&isnumber=11544166

[24] Y. Mittra, "Using Twitter to Measure the Impact of Immigration by Studying People's Mood," in *Proc. IEEE*, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9635537/

[25] S. Sekhri, "Advancements in Nucleation Theory: Extensions to the XPFC Model and Applications in Precipitation," 2025. doi: https://doi.org/10.5281/zenodo.17627211