dlweek

2025

MLDA@EEE

DATATHON

Content.

Background

Every year, thousands of expectant mothers undergo monitoring during labor using a tool called cardiotocography (CTG). CTG records both the baby's heart rate and the mother's uterine contractions, producing signals that clinicians interpret to detect early signs of fetal distress.

When recognized in time, doctors and midwives can intervene to prevent complications. However, CTG readings are complex, patterns can be subtle, and busy hospital wards make it difficult to spot problems quickly. This challenge asks you to reimagine how data can support timely, life-saving decisions.

Objective

The main objective of this challenge is to use real-world CTG recordings to build a solution that supports clinicians in spotting fetal distress. By analyzing patterns in the data, you are expected to design a system that can reliably separate Normal, Suspect, and Pathologic cases.

In doing so, your solution should not only focus on achieving good performance but also show that you understand the medical significance of the features and can present your findings in a way that makes sense to both technical and non-technical audiences.

Problem Statement

How can we build a solution that carefully **interprets the patterns hidden** in a **baby's heart rate** and the **mother's contractions**, in order to automatically **identify signs of fetal distress** that might otherwise go unnoticed during labor?

Expected Tasks

01. Data Analysis & Processing

Teams start by exploring the dataset and learning what each medical feature means (e.g. accelerations, decelerations, variability). They must clean errors, handle missing values, and prepare the data for analysis — building both medical understanding and computational readiness.

02. AI Model Development & Performance Evaluation

Next, teams design models to classify fetal states as Normal, Suspect, or Pathologic. They should track performance, test reliability, and reflect on fairness, ensuring their solution is useful beyond just raw accuracy.

03. Final Presentation & Communication

Finally, teams present their work through a slide deck, explaining methods, findings, and real-world implications. Clear storytelling and honest reflection on limitations will matter as much as technical results.

General Flow (Current Idea):
Round 1: general competition (leaderboard based).
Round 2: Top six to present their solution.

Familiarizing Yourself with the Cardiotocography

When a baby is still inside the womb, the only way we know if they're doing well is by listening to how their body is responding to stress. The main signals doctors look at are the baby's heart rate and the mother's contractions.

These two signals tell a story: imagine watching a graph of the baby's pulse while the womb tightens around them — does the baby recover quickly? Do they struggle? Do their heartbeats stay steady or fluctuate too much?

That's exactly what Cardiotocography (CTG) records. It's a non-invasive tool, meaning no needles or blood tests; instead, sensors are placed on the mother's abdomen. The machine prints two squiggly lines:

01. **Fetal Heart Rate (FHR)**
The rhythm of the baby's heartbeat, measured in beats per minute. Think of this like a continuous ECG, but for a fetus.

02. **Uterine Contractions (UC)**
The mother's tightening and relaxing of the womb, which squeezes blood supply temporarily.

Now here's the tricky part: doctors don't just look at one number (say, heart rate = 140 bpm). They watch for patterns over time. A healthy baby's heart should be like a lively conversation, sometimes speeding up, sometimes slowing down, always showing some variability. A distressed baby, in contrast, may have a heart rhythm that's too flat, too erratic, or too slow to bounce back after stress.

The challenge? Real wards are noisy, crowded, and often understaffed. A doctor may be looking at several CTG monitors at once, especially during the night. Subtle warning signs, like a delayed dip in heart rate after contractions, can be easily missed by the human eye. That's why this datathon exists: to explore whether AI can become a second pair of eyes, catching patterns that slip through in real time.

Understanding the Dataset

To work on this problem, we'll use the UCI Cardiotocography (CTG) dataset, a widely studied collection of recordings.

Other datasets are allowed, however, we recommend this specific dataset. Over 2,000 real CTG traces were reviewed by three obstetricians, and a consensus label was assigned: Normal, Suspect, or Pathologic. This makes the dataset a trustworthy ground truth for us.

Let’s break down some of the most important features — and I’ll explain them as if you’re sitting in a delivery ward, watching the monitor.

LB (Baseline Heart Rate): Imagine the baby’s “resting pulse” over one minute. A healthy baseline is usually 110–160 bpm. Too low (bradycardia) might mean the baby isn’t getting enough oxygen. Too high (tachycardia) could signal infection, maternal fever, or fetal stress.

AC (Accelerations): These are short, reassuring “spikes” where the baby’s heart beats faster for a moment (like your pulse jumping after a sudden laugh). Frequent accelerations usually mean the baby is active and oxygenated.

FM (Fetal Movements): These are the kicks or rolls detected. Just like you’d worry if someone sat unnaturally still for hours, doctors worry if movements are absent for too long.

UC (Uterine Contractions): Every time the womb contracts, it briefly squeezes blood vessels. A healthy baby tolerates these dips without problem. A distressed baby’s heart may dip sharply during or after contractions.

DL (Light Decelerations): Small, shallow dips in heart rate. They’re not always dangerous — think of them as little “hiccups” in the rhythm.

DS (Severe Decelerations): Sharp, deep drops, like the baby suddenly losing oxygen. These are red flags.

DP (Prolonged Decelerations): Heart rate doesn’t bounce back quickly — it stays low for too long. This is particularly worrying and often requires intervention.

ASTV (Abnormal Short-Term Variability): Variability means “wiggliness.” If the heart rate line is completely flat, it’s like a monotone voice — concerning because it suggests the baby isn’t responding well. Too much erratic movement, on the other hand, can also be abnormal.

ALTV (Abnormal Long-Term Variability): Looks at the rhythm over longer stretches, not just second-to-second changes. Again, too little or too much can be problematic.

Histogram Features (Width, Min, Max, Mode, Median, Variance, etc.): These are statistical summaries of the heart rate distribution. Imagine we take the one-minute heart rate trace and turn it into a histogram.

a. *Width tells us how spread out the heart rates are (lots of ups and downs vs. mostly steady).*
b. *Min and Max show extremes (lowest and highest bpm).*
c. *Variance tells us how much the baby’s heart rate changes moment to moment.*

Together, these features form a kind of fingerprint of fetal well-being. A healthy CTG usually shows:

- A stable baseline,
- Some accelerations,
- Variability that is “lively” but not chaotic, and few or no severe decelerations.

Whereas a pathological CTG may show:

- A too-low or too-high baseline,
- Repeated deep decelerations,
- Flat or abnormal variability, and little movement or accelerations.

The dataset labels each case as:

- **Normal (N):** Baby is well, no immediate concern.
- **Suspect (S):** Warning signs present, needs careful monitoring.
- **Pathologic (P):** High risk of distress urgent attention required.

This is why the dataset is so valuable: it doesn’t just give us numbers, it gives us clinical meaning behind those numbers.

Getting Started with Data

The very first step in solving any real-world problem is understanding the data you're working with. In medicine, data isn't just numbers — it's evidence of how a patient's body is functioning. In our case, each row in the dataset is a snapshot of how a baby's heart was behaving during pregnancy, as interpreted through a cardiotocogram (CTG).

Think of it like this:

Each feature (LB, AC, FM, etc.) is a clue. Together, all the features form a case file. Your job is to sift through these case files and figure out whether the baby looked healthy, borderline, or in trouble.

Processing the Data

When you first open the dataset, it will look like a spreadsheet full of numbers and abbreviations. Doctors can read these easily because they know the medical context, but to a computer, “LB” or “ASTV” means nothing. You need to translate the medical language into something a machine can understand.

Here’s how to think about it:

Cleaning the data:
You need to make sure your dataset isn’t full of errors or missing entries. A blank cell in a spreadsheet is like a missing heartbeat on a monitor which can confuse your model if not handled carefully.

Scaling and normalizing:
Imagine you’re comparing two patients: one has a baseline of 120 bpm, another 160 bpm. If another feature (say, accelerations) is measured on a much smaller scale, the model might “ignore” it unless you bring everything to the same footing. That’s what scaling does: it ensures all features speak with the same voice.

Encoding the labels:
The medical experts gave us categories like Normal, Suspect, Pathologic. Computers, however, don’t understand words, only numbers. So you’ll map these to codes (for example, Normal = 0, Suspect = 1, Pathologic = 2). This doesn’t change the meaning — it just makes it machine-readable.

There are various other ways to make the data readable by your Machine Learning model so feel free to explore and research as you like!

Exploring Relationships

Before rushing into building an AI model, pause and listen to what the data is telling you. Doctors often notice patterns like: Babies with high ASTV tend to be riskier. Frequent severe decelerations almost always point to pathology. You can uncover similar insights by using tools like:

Heatmaps

These visualize correlations (relationships) between features. For instance, you might discover that accelerations and fetal movements often rise together, while severe decelerations strongly link with pathologic labels.

Histograms and boxplots

These let you peek at the spread of heart rate values, just like a doctor glancing at a CTG printout.

The point isn't just to draw pretty graphs — it's to develop intuition. This is why Data Exploration is important. You need to understand and find patterns in the dataset. A good competitor in this datathon isn't just someone who codes fast, but someone who can say:

> "I noticed that babies with very low variability almost always ended up in the Pathologic group."

That's the level of reasoning we're looking for.

Building the Model

This is the exciting part: you get to teach a computer how to “read” CTGs like a junior doctor. There’s no single “right” algorithm, but here’s how you might think about it:

Start simple

Logistic regression or decision trees are like the stethoscope of machine learning — basic, but surprisingly powerful. They let you see which features matter most.

Go deeper

Random forests or gradient boosting methods combine many decision trees to capture complex interactions (e.g., “if baseline is high and ASTV is abnormal, risk is higher”).

Experiment

Try multiple models. Some may overfit (memorize the training set instead of generalizing), while others may underfit (too simple to catch the patterns). The trick is to balance both.

Remember

we’re not just asking for the most accurate model. We’re asking for a solution that makes sense clinically. A flashy deep learning model with 98% accuracy is less impressive if you can’t explain which features drove the decision.

Evaluating Performance

We will judge models based on:

Balanced Accuracy

Ensures that all three classes (Normal, Suspect, Pathologic) are treated fairly, even if some have fewer samples.

Macro F1-Score

A measure that balances precision and recall across all classes. It answers: How good is your model at catching each condition without overcalling?

Judging Criteria:

- Problem framing & data hygiene (20%)
- Imbalance handling & metrics literacy (20%)
- Model quality (Balanced Acc, Macro F1) (30%)
- Explainability & clinical reasoning (15%)
- Latency & engineering (15%)

GENERAL HACKATHON SCHEDULE

Event	Start	End	Duration
Opening Ceremony	29 Sep, 20:00	29 Sep, 21:00	1 hour
Build Sprint	29 Sep, 20:00	5 Oct, 23:59	6 days
Top 10 Announced	6 Oct, 23:59	6 Oct, 23:59	NA
Top 10 Preparation for Demo Day	7 Oct, 0:00	10 Oct, 16:30	3 days
Demo Day	10 Oct, 16:30	10 Oct, 21:00	5 hours

CORE CONCEPTS

Basic topics

1. What a classifier is (mapping inputs → categories).
2. How to split data into train/test sets (bonus: stratified splitting for class balance).
3. Class imbalance basics (why accuracy is misleading, why minority classes get ignored).
4. Balanced Accuracy = average recall across classes.
5. RandomForest / Logistic Regression → how they work at a very high level (trees combine votes, logistic regression fits probabilities).

Advanced topics (optional but may help)

1. Handling class imbalance (SMOTE, class weights, oversampling/undersampling).
2. Neural networks basics (MLPs, CNNs, RNNs for sequences).
3. Regularization (dropout, L2).
4. Model interpretability (feature importance, SHAP/LIME basics).
5. Model evaluation beyond accuracy (macro-F1, ROC curves).
6. Efficiency constraints (lightweight architectures, avoiding overfitting).

Study Materials

- Scikit-learn Beginner Tutorial
- Train/test split & evaluation metrics
- Balanced Accuracy explainer
- Imbalanced-learn documentation
- **Interpretable Machine Learning**
  by Christoph Molnar (free online)
- Scikit-learn MLPClassifier
- Intro to SHAP

Checklist

- You're good to go if you can:
- Train a RandomForest or Logistic Regression.
- Get predictions for 3 classes.
- Report Balanced Accuracy + Macro-F1.
- Experiment with neural nets. (optional)
- Justify your model/tradeoffs in plain English.

Sign up to get the latest updates on partner company events, resume-boosting opportunities, internship fast-track programs and more MLDA-affiliated events.

Membership Form:
https://tinyurl.com/dlwmembership

Event Feedback Form:
https://tinyurl.com/dlweventfeedback

Datathon Telegram Channel:
https://tinyurl.com/datathon2025telegram