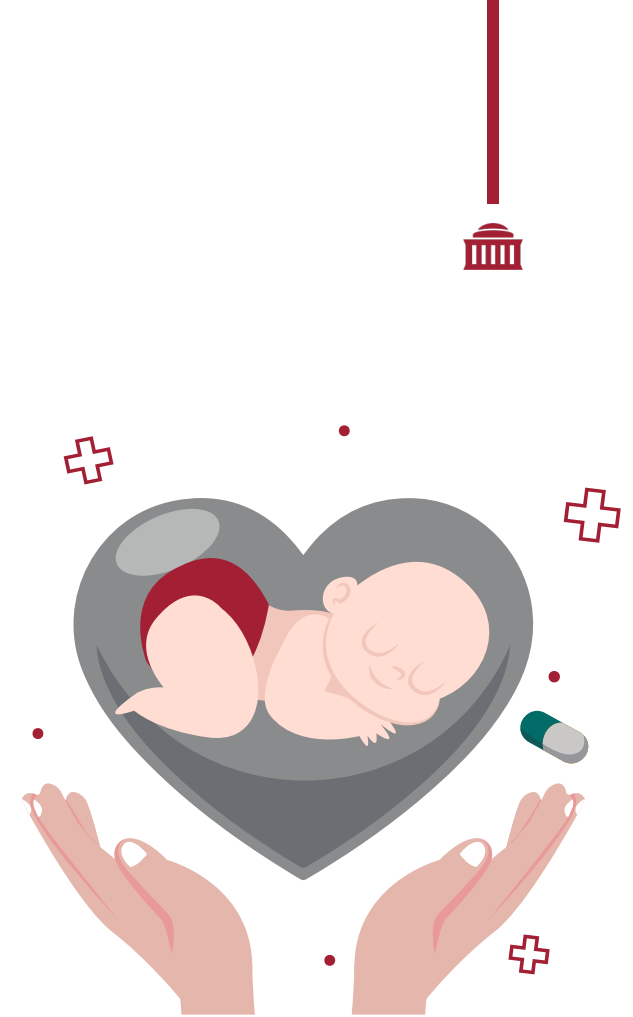


# EMPOWERING PRENATAL DIAGNOSTICS & CARE

Multiclassification of fetal health status using CTG  
and interpretable clustering for diagnosis

Angel Gantzia  
Aditi Singh

MIT SLOAN SCHOOL  
OF MANAGEMENT



# PROBLEM WITH CURRENT FETAL MONITORING PRACTICES



**“One death every 7 seconds** among women and babies during pregnancy and childbirth, mostly from **preventable** or **treatable** causes”



Cardiotocography (CTG) is a non-invasive fetal surveillance method



Complex CTG patterns associated with fetal hypoxia (common risk factor) are generally assessed visually; results poorly reproducible and inconsistent

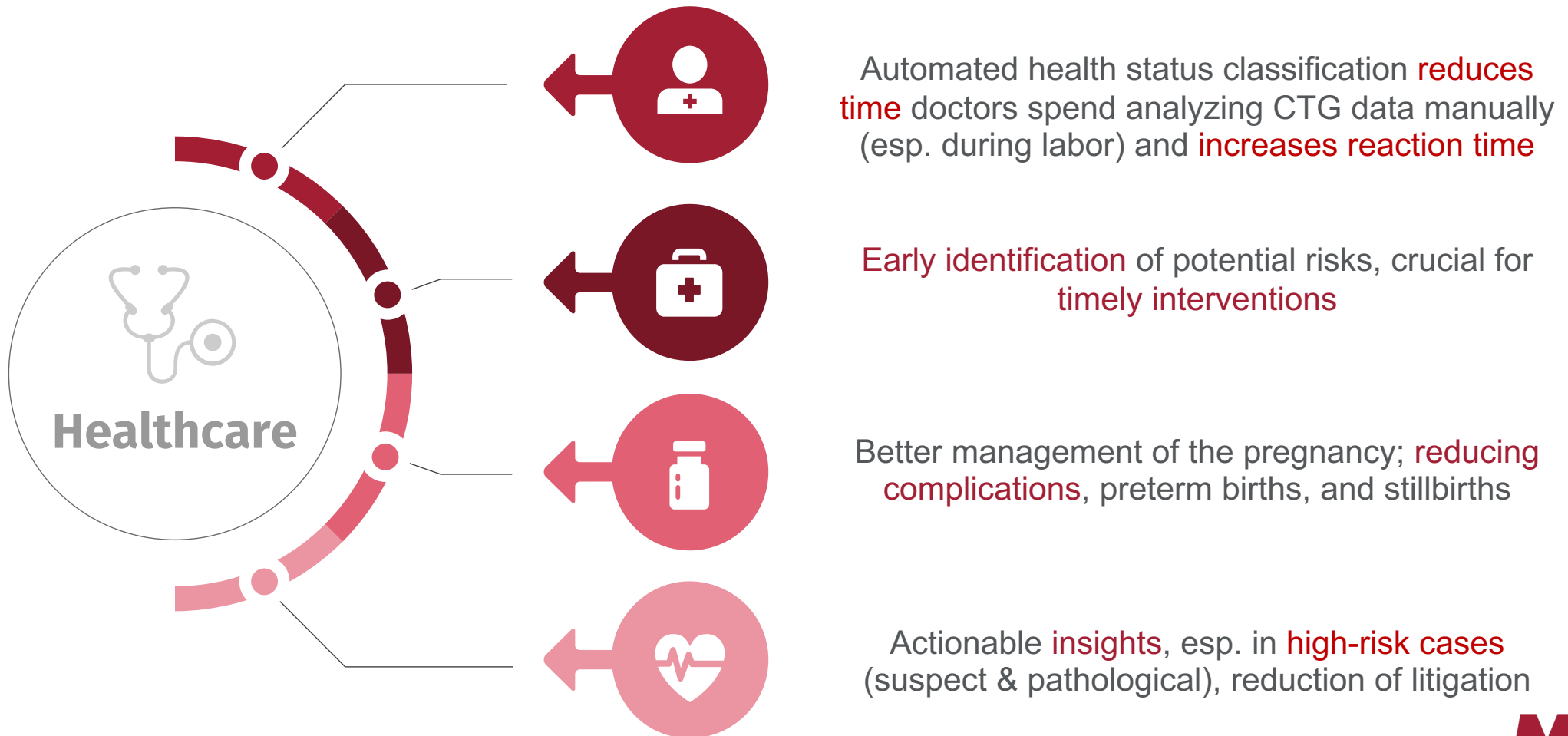


Automated CTG Classification could help optimize the diagnostic accuracy



# BENEFITS

## SAFER PREGNANCIES AND HEALTHIER BABIES



# DATA

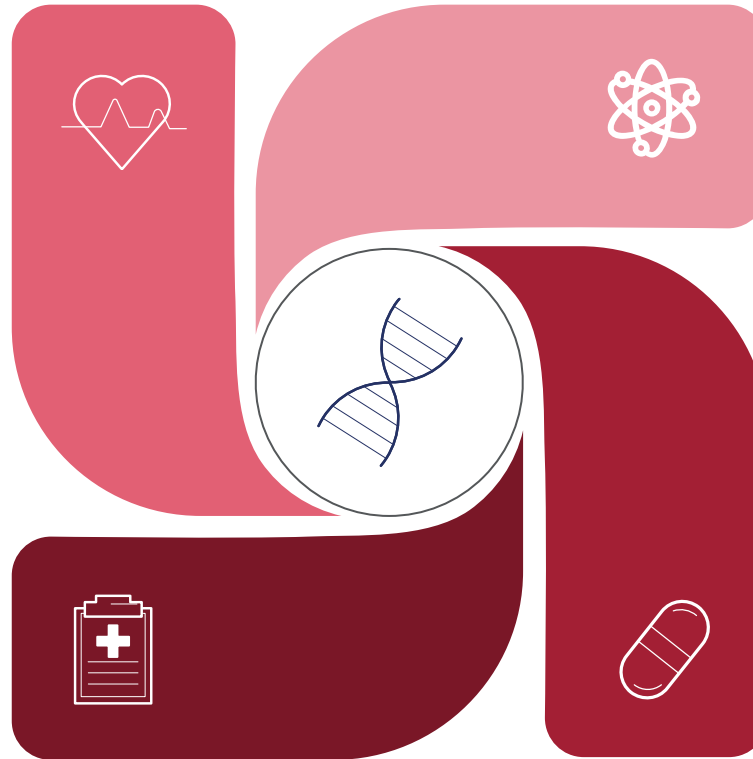
## IMBALANCED DATA, 2126 CASES WITH 21 PREDICTORS

### Heart Rate

Baseline  
Accelerations  
Decelerations (# of  
light, severe, prolonged)

### Fetal Movement and Uterine Activity

# of fetal Movements  
# of uterine Contractions



### Variability

% and mean of abnormal  
short-term Variability  
% and mean of abnormal  
long-term Variability

### Histogram Characteristics

Histogram width, min, max  
# of peaks and zeroes  
Mode, mean, median  
Variance, tendency

# METHODOLOGY

## 5 Tuned Classifiers

with cross validation  
using average F1-Score



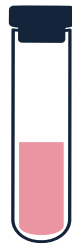
## Interpretable Clustering

with cross validation  
using average F1-Score

## Other Methods & Experiments

### Data Standardization

Different scales or units  
(e.g., heart rate in bpm, number of  
movements per second)



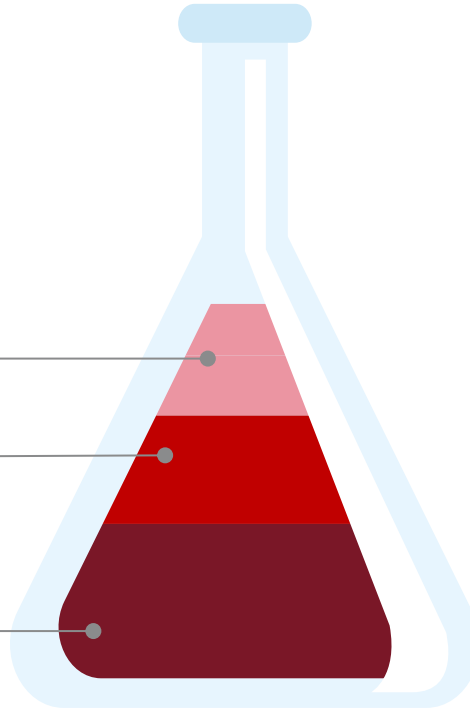
### Non-linear transformations

Multiplying features  
& Individual transformations  
(squared, square roots)



### SMOTE

Synthetic Minority Oversampling  
to balance the training set



Removing /  
Swapping out highly  
correlated predictors

Consulting with resident doctor for  
interpretability and clinical relevance  
of clusters and data



Applying Soft-Voting  
to improve  
performance

Applying different  
hyperparameters and  
scoring metrics



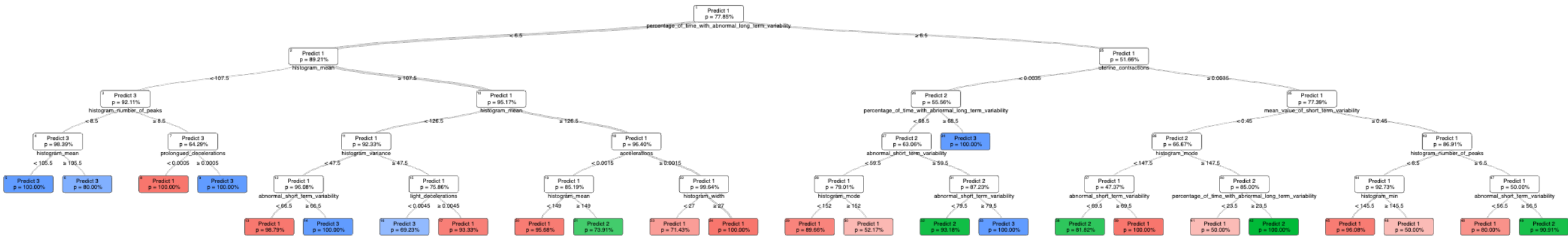
# KEY RESULTS



Model	Scaled	SMOTE	Non-linear Transformations	Accuracy	Weighted F1 Score	Weighted Precision	Weighted Recall
Support Vector Classifier	✓		✓	92.48%	92.30%	92.22%	92.49%
Random Forest Classifier		✓	✓	91.54%	91.83%	92.34%	91.55%
OCT				89.10%	89.73%	91.36%	89.10%
CART		✓	✓	89.10%	89.39%	89.84%	89.10%
Logistic Regression	✓	✓	✓	88.35%	89.20%	91.02%	88.35%

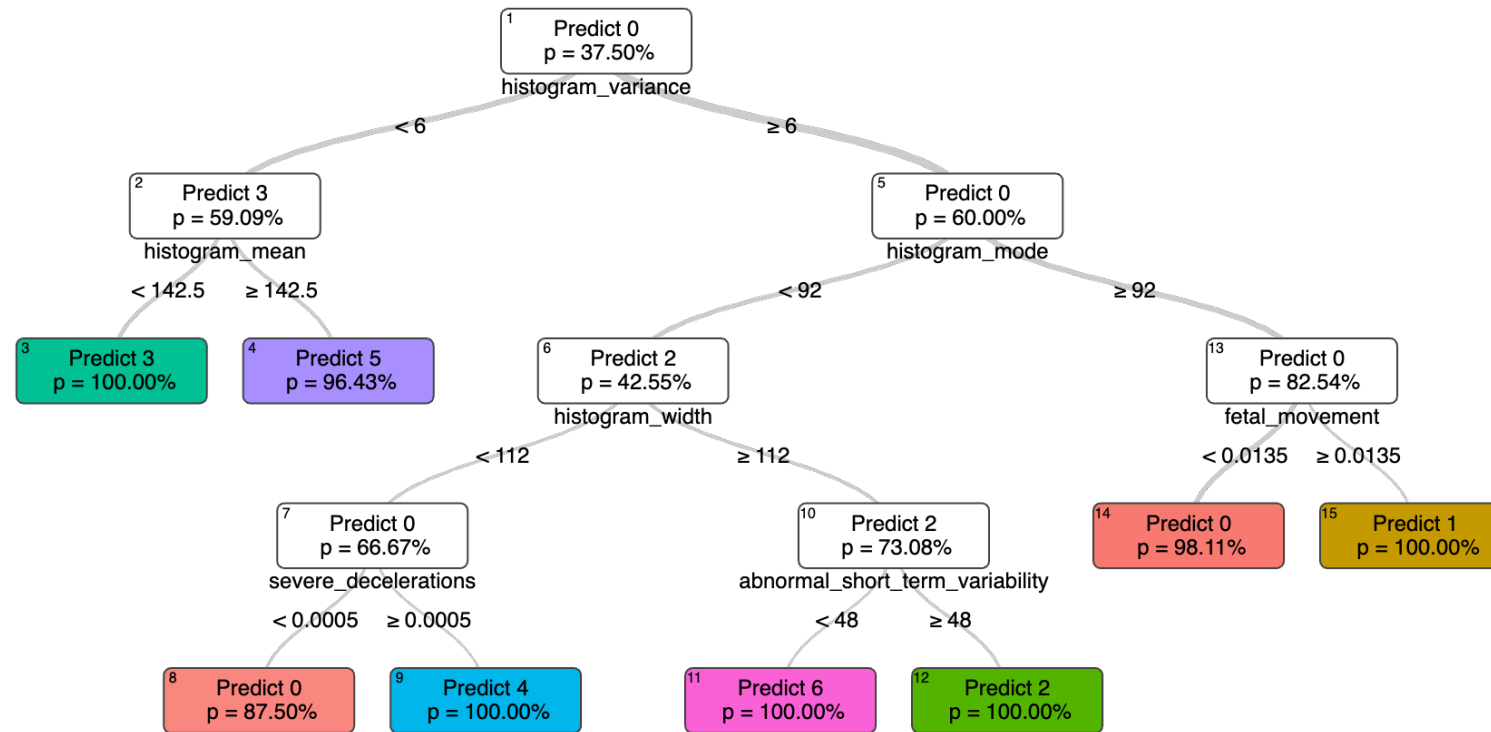
- Best models in each category, sorted by weighted F1-Score, given that all individual scores are above 70%
- Vast improvement over Baseline Logistic Regression, with 85% weighted F1, individual scores as low as 40%
- In the accuracy trade-off, we recommend the use of OCT, which has the sizable benefit of interpretability
- Soft voting improves accuracy further (93%), but loses on the interpretability

# HOW DO WE TO DIAGNOSE?



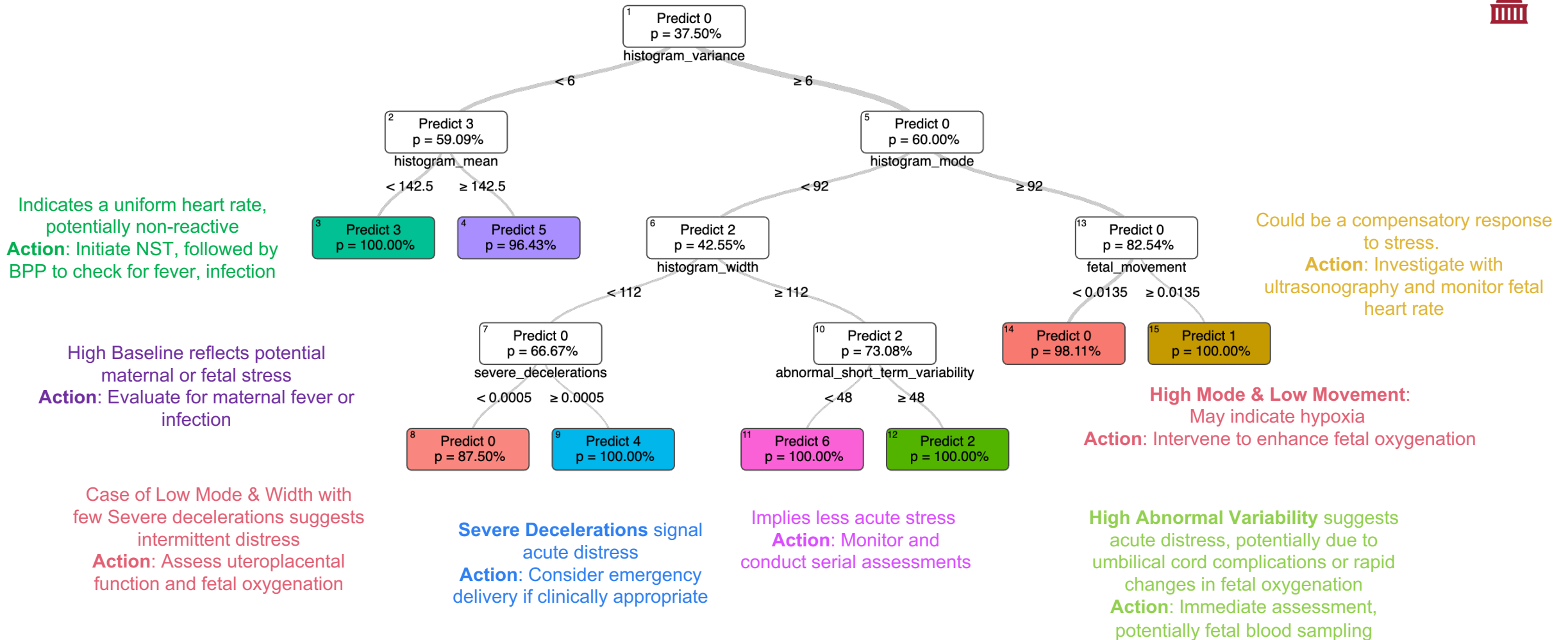
But **what actions** can we take if the fetus is diagnosed as Suspect or Pathological?  
**Why? How?**

# ACTIONABLE INSIGHTS FOR PATHOLOGICAL CASES





# ACTIONABLE INSIGHTS FOR PATHOLOGICAL CASES



# THANK YOU

Do you have any questions?