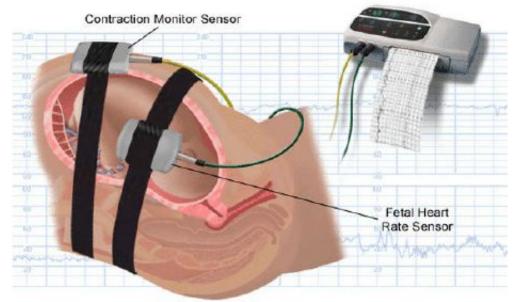
Classification of anomalous fetal status from cardiotocographies

Ilaria Bonavita

Cardiotocography

- Method for monitoring the fetal heart rate (FHR) e the uterine contractions (UC)
- The analysis of a cardiotocography (CTG) and of the features extracted from it, allow to assess the health status of the fetus



Non-Invasive Fetal Heart Rate Monitoring Techniques. Available from: https://www.researchgate.net/figure/An-external-fetal-cardiotocography-CTG-monitor_fig9_268515126

The problem

Given a set of features extracted from cardiotocographies, create a model of health status of the fetus (normal, suspect or pathologic)

Motivation

- Some features are difficult to be analysed by human eye
- Visual analysis of CTGs has proven low consistency
- Automatised tools can help the experts in the more ambiguous cases
- Optimisation in the usage of human and technological resources of the hospitals

Workflow

Problem **Understanding** (bibliographic research, etc.)

Data preprocessing

Explorative analysis (visualisation, descriptive statistics)

Modelling (parameter tuning, validation)

Results interpretation (evaluation metrics)

The dataset

Ayres de Campos et al. (2000), UCL Machine Learning repository

2126 samples

CTGs collected between 1980-98 by Porto Universitary Hospital

21 features

Extracted and processed from FHR according to medical guidelines

3 classes

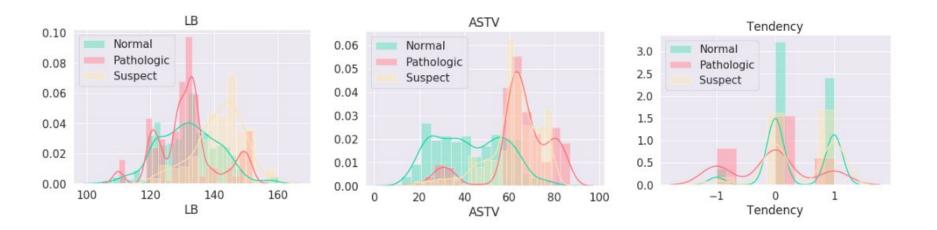
obtained with majority vote by 3 experts obstetricians

LB	Baseline FHR			
FM	Fetal movement			
UC	Uterine Contractions	Class	N samples	Percentage
AC, DL, DS, DP	Accelerations,	Normal	1655	77.8 %
	light/severe/prolonged decelerations	Suspect	295	13.9 %
ASTV, MSTV, ALTV, MLTV	Short and long term variability stats	Pathological	176	8.3 %
Width, Min, Max, Nmax, Mean, Mode, Median, Variance, Tendency, Nzeros	FHR Histogram stats			

Features

Meaning

Explorative Analysis



LB	1	-0.078	-0.033	-0.13	-0.16	-0.054	-0.12	0.31	-0.28	0.29	-0.032	-0.15	0.36	0.28	-0.11	-0.0047	0.71	0.72	0.79	-0.13	0.29	-0.25	-0.049	0.34
AC	-0.078	1	0.054	0.21	-0.011	-0.039	0.11	-0.2	0.19	-0.35	-0.13	0.29	0.16	0.37	0.19	-0.011	0.22	0.24	0.25	0.11	0.032	0.37	-0.2	-0.28
FM	-0.033	0.054	1	-0.037	0.076	-0.01	0.21	-0.12	0.14	-0.072	0.021	0.17	-0.16	0.11	0.18	-0.015	-0.037	-0.086	-0.056	0.18	-6.9e-05	-0.044	0.068	-0.0017
UC	-0.13	0.21	-0.037	1	0.34	0.018	0.15	-0.08	0.23	-0.25	-0.11	0.15	-0.11	0.15	0.096	0.023	-0.078	-0.17	-0.11	0.18	-0.061	0.21	-0.042	-0.22
DL	-0.16	-0.011	0.076	0.34	1	0.12	0.29	-0.037	0.52	-0.27	-0.25	0.52		0.24	0.39	0.21	-0.32		-0.35	0.52	0.00069	0.03	0.19	0.19
DS	-0.054	-0.039	-0.01	0.018	0.12	1	0.02	0.034	0.034	-0.031	-0.038	0.045	-0.072	-0.021	0.007	0.043	-0.22	-0.16	-0.16	0.14	-0.07	-0.088	0.16	-0.023
DP	-0.12	-0.11	0.21	0.15	0.29	0.02	1	0.054	0.29	-0.14	-0.23	0.28	-0.29	0.13	0.24	0.046	-0.41		-0.45	0.49	-0.21	-0.34	0.58	-0.053
ASTV	0.31	-0.2	-0.12	-0.08	-0.037	0.034	0.054	1	-0.43	0.46	-0.32	-0.26	0.28	-0.11	-0.17	-0.15	0.058	0.075	0.12	-0.15	-0.0057		0.31	0.35
MSTV	-0.28	0.19	0.14	0.23	0.52	0.034	0.29	-0.43	1	-0.47	0.074		-0.62	0.41	0.5	0.27	-0.31	-0.45	-0.34	0.56	-0.066	0.21	0.083	-0.32
ALTV	0.29	-0.35	-0.072	-0.25	-0.27	-0.031	-0.14	0.46	-0.47		-0.17	-0.45	0.42	-0.28	-0.28	-0.12	0.17	0.22	0.19	-0.28	0.042		0.21	0.42
MLTV	-0.032	-0.13	0.021	-0.11	-0.25	-0.038	-0.23	-0.32	0.074	-0.17	1	0.11	-0.14	0.002	0.056	0.12	0.072	0.14	0.063	0.16	0.15	0.17	-0.25	-0.012
Width	-0.15	0.29	0.17	0.15	0.52	0.045	0.28	-0.26		-0.45	0.11					0.32	-0.16	-0.28	-0.17		0.12	0.14	0.061	-0.22
Min	0.36	-0.16	-0.16	-0.11		-0.072	-0.29	0.28		0.42	-0.14			-0.3		-0.31	0.35	0.49	0.4		-0.24	-0.16	-0.098	0.27
Max	0.28	0.37	0.11	0.15	0.24	-0.021	0.13	-0.11	0.41	-0.28	0.002		-0.3		0.52	0.18	0.24	0.19	0.29	0.44	-0.14	0.048	-0.028	-0.035
Nmax	-0.11	0.19	0.18	0.096	0.39	0.007	0.24	-0.17	0.5	-0.28	0.056			0.52		0.29	-0.1	-0.22	-0.12	0.45	0.11	0.06	0.038	-0.1
Nzeros	-0.0047	-0.011	-0.015	0.023	0.21	0.043	0.046	-0.15	0.27	-0.12	0.12	0.32	-0.31	0.18	0.29	1	-0.058	-0.084	-0.053	0.2	0.085	0.031	0.0098	-0.045
Mode	0.71	0.22	-0.037	-0.078	-0.32	-0.22	-0.41	0.058	-0.31	0.17	0.072	-0.16	0.35	0.24	-0.1	-0.058				0.31	0.42	0.092	-0.42	0.22
Mean	0.72	0.24	-0.086	-0.17		-0.16		0.075	-0.45	0.22	0.14	-0.28	0.49	0.19	-0.22	-0.084				-0.4	0.33	0.059	-0.42	0.26
Median	0.79	0.25	-0.056	-0.11	-0.35	-0.16	-0.45	0.12	-0.34	0.19	0.063	-0.17	0.4	0.29	-0.12	-0.053				-0.29	0.39	0.048	-0.38	0.25
Variance	-0.13	0.11	0.18	0.18	0.52	0.14	0.49	-0.15	0.56	-0.28	-0.16			0.44	0.45	0.2	-0.31	-0.4	-0.29		-0.078	-0.086	0.33	-0.16
Tendency	0.29	0.032	-6.9e-05	-0.061	0.00069	-0.07	-0.21	-0.0057	-0.066	0.042	0.15	0.12	-0.24	-0.14	0.11	0.085	0.42	0.33	0.39	-0.078	1	0.067	-0.19	0.075
class_Normal	-0.25	0.37	-0.044	0.21	0.03	-0.088	-0.34		0.21		0.17	0.14	-0.16	0.048	0.06	0.031	0.092	0.059	0.048	-0.086	0.067			0.75
class_Pathologic	-0.049	-0.2	0.068	-0.042	0.19	0.16	0.58	0.31	0.083	0.21	-0.25	0.061	-0.098	-0.028	0.038	0.0098	-0.42	-0.42	-0.38	0.33	-0.19		1	-0.12
class_Suspect	0.34	-0.28	-0.0017	-0.22	-0.19	-0.023	-0.053	0.35	-0.32	0.42	-0.012	-0.22	0.27	-0.035	-0.1	-0.045	0.22	0.26	0.25	0.16	0.075		-0.12	
	IB	AC	FM	OC	DF	DS	DP	ASTV	WSTV	ALTV	MLTV	Width	Min	Max	Nmax	Nzeros	Mode	Mean	Median	Variance	Tendency	class_Normal	ass_Pathologic	class_Suspect

-0.8

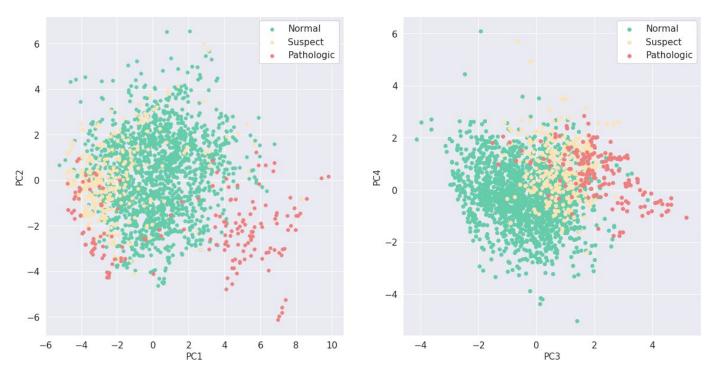
-0.4

-0.0

--0.4

- -0.8

Principal Components, dimensions 1 to 4



4 PCs -> 60 % variability 11 PCs -> 90% variability

Useful observations for choosing the model

- Highly imbalanced classes
- Presence of highly correlated features
- Heterogeneous features (different ranges, type and distribution)
- N samples >> N features
- Non-trivial separation among classes

Random Forest Classifier

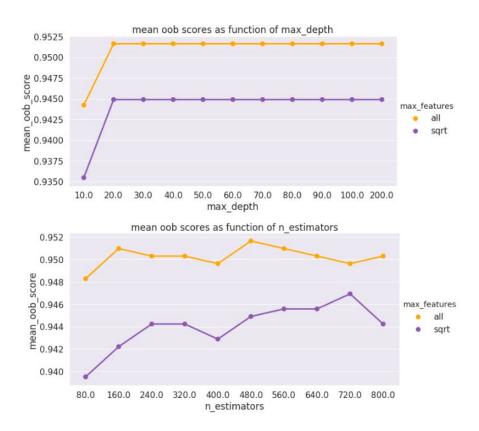
- Dataset annotated through a decision process
- + Suitable for a wide range of problems
- + **Robust** with respect to outliers, collinearity and heterogenous variables
- + Does not require statistical hypothesis on the data
- + Suitable for cases of class imbalance
- Computational times
- Hypertuning of the parameters usually needed

Training

- Train test 70-30 split
- Grid search for hyperparameters tuning
 - N estimators, max depth, min samples leaf, max features
 - 660 configurations
 - ~ 19 minutes,
- Out-of-bag (OOB) score: 0.95161

Final model

Value
Gini
480
20
all
2
balanced
True



Results

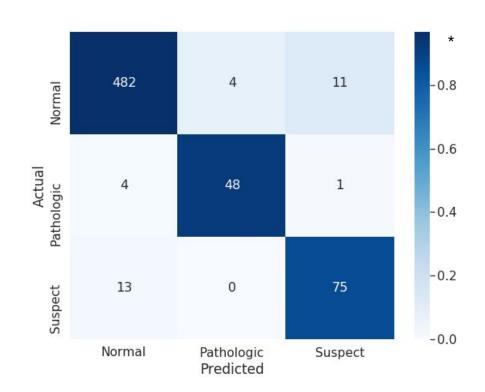
95 % 91 %

Accuracy F1-score*

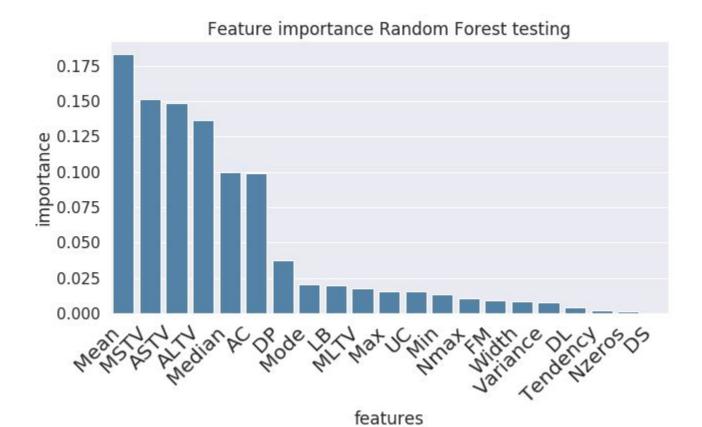
91 % 92 %

Precision* Recall*

	Precision	Recall	F1-score
Normal	0.97	0.97	0.97
Pathologic	0.91	0.92	0.91
Suspect	0.85	0.86	0.86



* the color refers to the normalised values per predicted



Results on a reduced features set

- 7 most important features
- Random forest with optimal obtained parameters
- Prediction on the same Test set

94% Accuracy

90% F1-score

Future steps

- Systematic analysis of the **stability** of the model (cross validation vs oob error, confidence intervals for metrics...) and of the **features importances**.
- Deeper inspection of the features of the misclassified **Suspect** samples
- Testing on at least another independent dataset
- Collecting feedback from obstetricians and medical experts
- Literature research and application of state-of-the-art models

Conclusions

- The model has a high predictive (91% F1-score) and generalization (00B score 95%) abilities
- 7 features are particularly informative and interpretation of FHR could focus on these
- The model could provide useful support in difficult clinical decisions
- CTGs classified as **suspect** require more in-depth analysis



