Metis Online Flex, Anu Adeboje

The Cesarean Decision

PREDICTION VIA CLASSIFICATION

"Rate Of C-Sections Is Rising, at an 'Alarming' Rate"

- FROM 6% IN 1990 TO 21% IN 2018, GLOBALLY
- C-SECTION CAN RAISE THE CHANCE OF DEATH BY AT LEAST 60%
- 3 FACTORS: FINANCIAL, LEGAL, AND TECHNICAL.

Objectives

Evidence Based Approach

Understand anthropometric, health, and obstetric attributes that influence the rise in cesarean section births.

Build a Classification Model

Leveraging identified attributes, predict whether a birth will be cesarean.

Model Interpretation

Identify drivers through feature analysis, identify next steps.

Larissa

Data & Methodology

DATA, CDC VITAL STATISTICS 2018

3,801,534 US Births

FEATURES EXTRACTED VARIABLES

anthropometric, health, and obstetric attributes

PREDICT, WILL BIRTH BE C-SECTION?

Build and test models, review feature importance

- **Cesarean section on a rise-Does advanced maternal age explain the increase? A population register-based study. Rydahl, E., Declercq, E., Juhl, M., & Maimburg, R. D. (2019).**

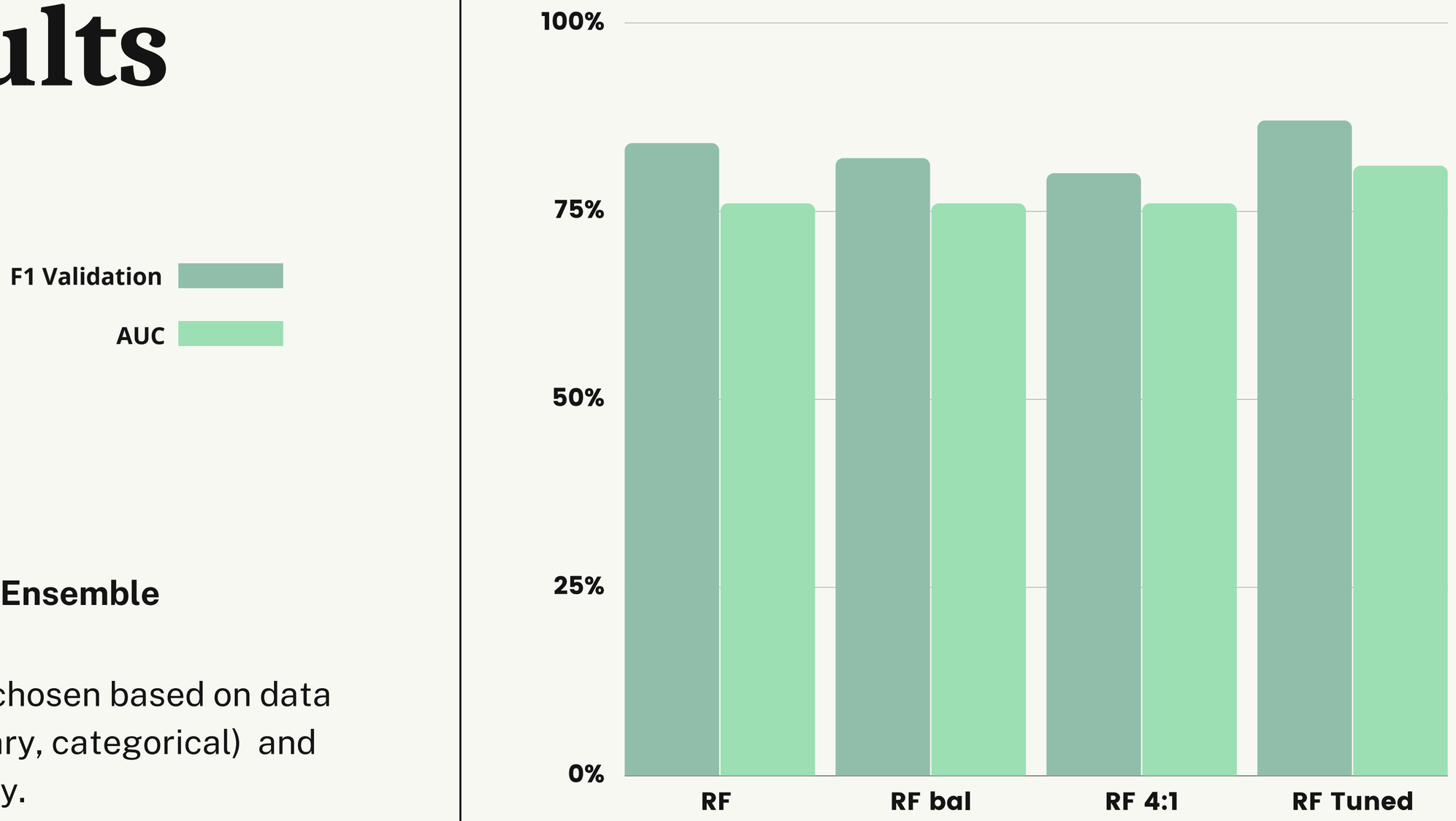
- **Improving the Caesarean Decision by Robson Classification: A Population-Based Study by 5,323,500 Livebirth Data.Tontus, H. O., & Nebioglu, S. (2020).**

Features

represent health features of the mother, reflect population trends, generally known prior to birth

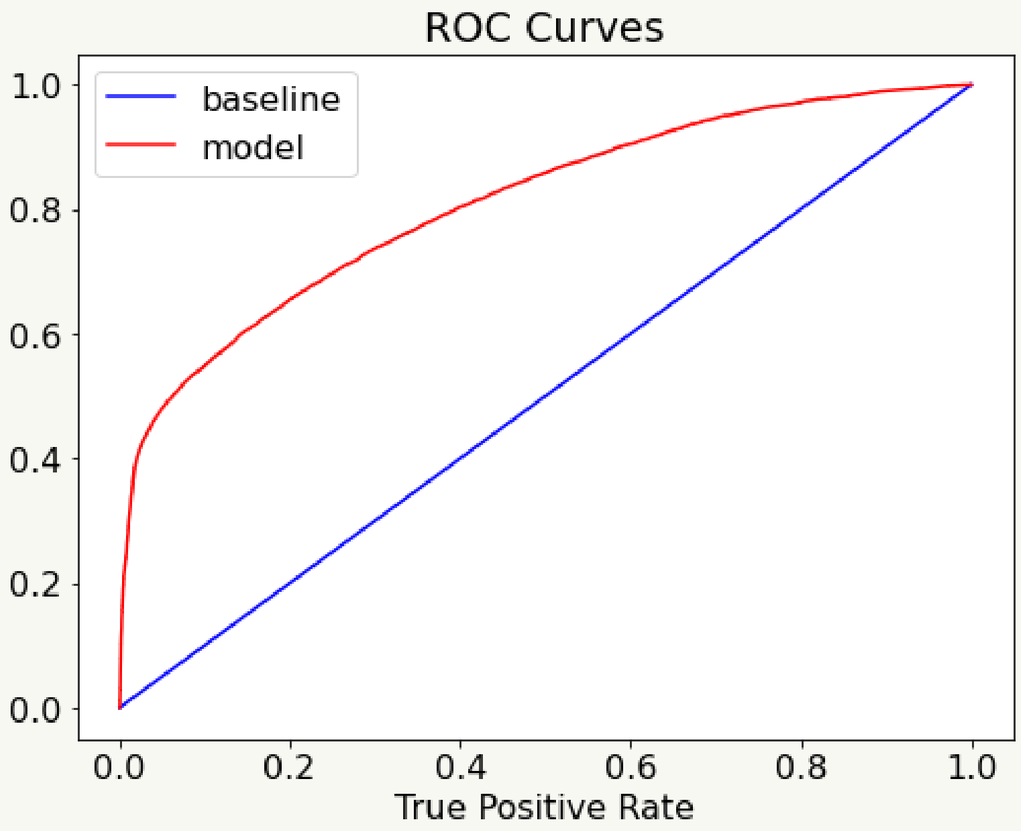
- Cigarettes Before Pregnancy
- Birth Month, & Day of Week
- Weight Gain by Mother
- Interval Since Last Pregnancy
- Live Birth Order/Total birth Order
- Induction of Labor
- Mother's Age
- Infections Reported
- Months of Prenatal Care
- Previous Cesarean
- Hypertension Eclampsia
- Infertility Treatment
- Prior Preterm Birth
- Sex of Infant

Results



Random Tree Ensemble

Models were chosen based on data structure(binary, categorical) and interpretability.



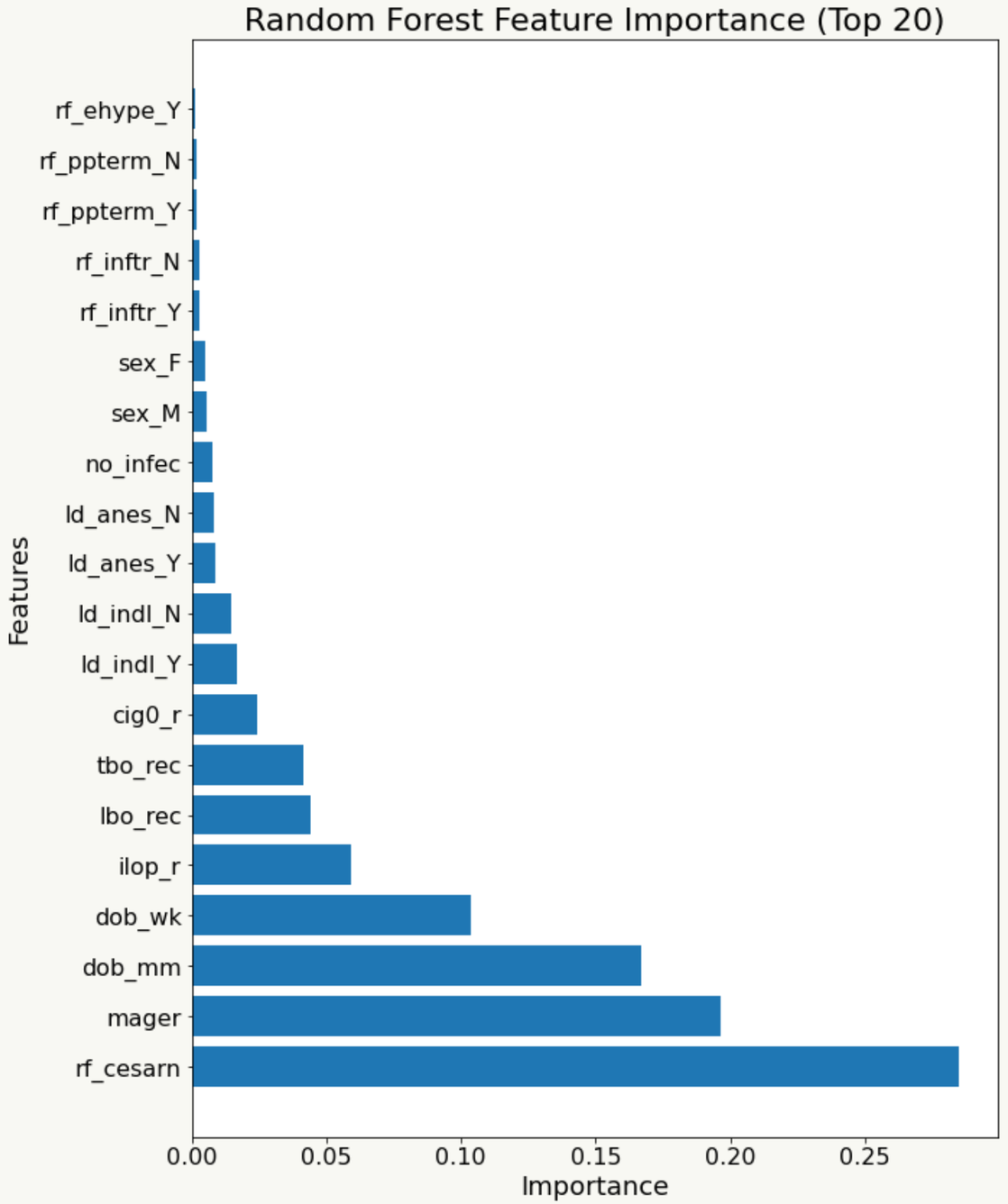
Results

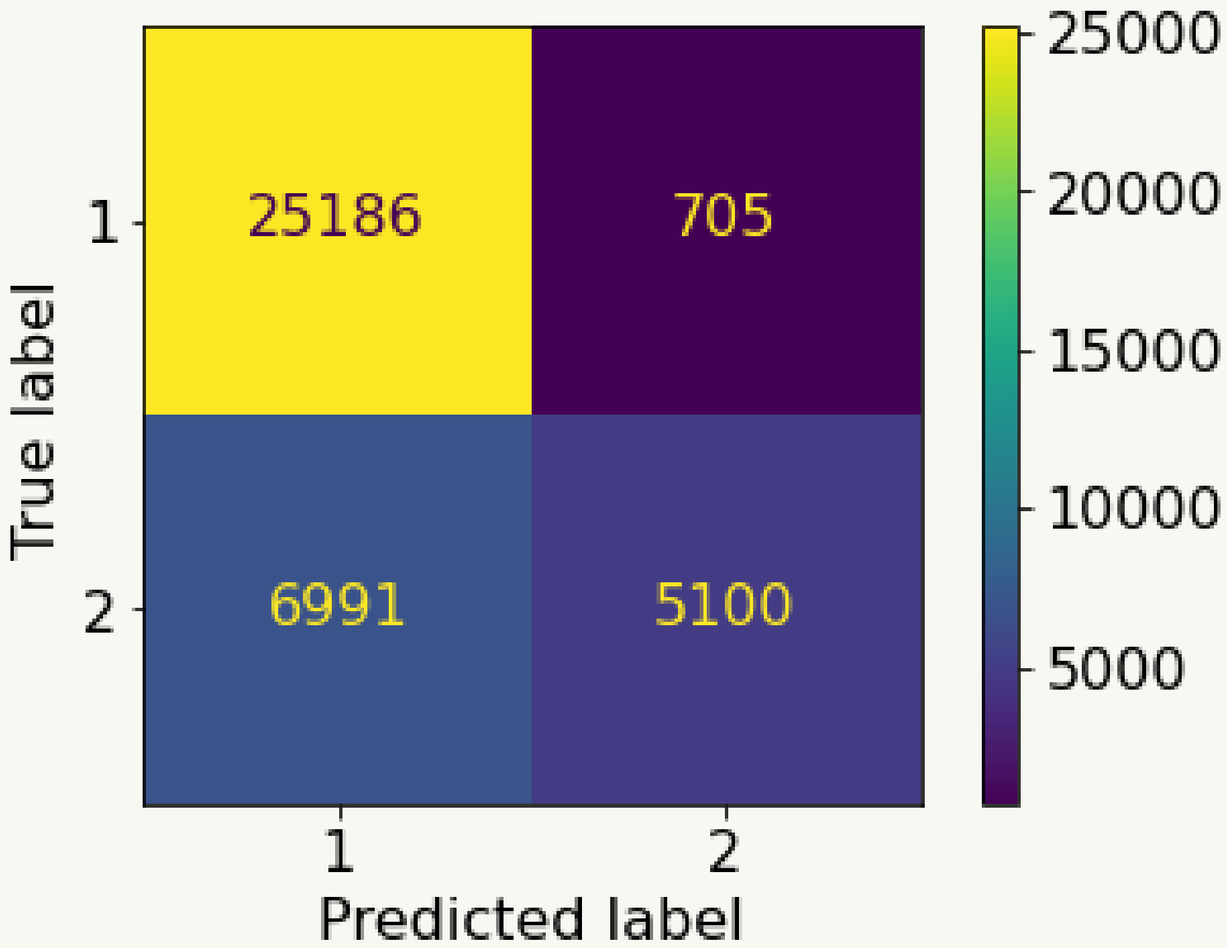
Best Model- Random Forest Tuned

After hyperparameter tuning we have an F1 score of 87% and ROC AUC score of 81% on the validation set, an improvement from our previous models. **The F1 for the test data set is: 87% and the AUC is 81%**

Feature Ranking

1. Prior C-section
2. Mother's Age
3. Month
4. Day of week
5. Interval since last Pregnancy





Conclusions

Our research suggested that the model might have a higher proportion of False Negatives, possibly due to elective or non health mandated C-sections (legal, financial)

- **False negatives = 6991, predicted vaginal birth**
- **False positives = 705**
- **True negatives = 25186**
- **True positives = 5100**

1-Vaginal Birth
2- C-section

Next Steps

- Apply model on test data from previous years to explore when the gap between health and elective c-sections occurs.
- Build regression models with infant outcomes, ie. Apgar score as targets to contrast feature importance in both models
- Summarize additional attributes of the False negative group.