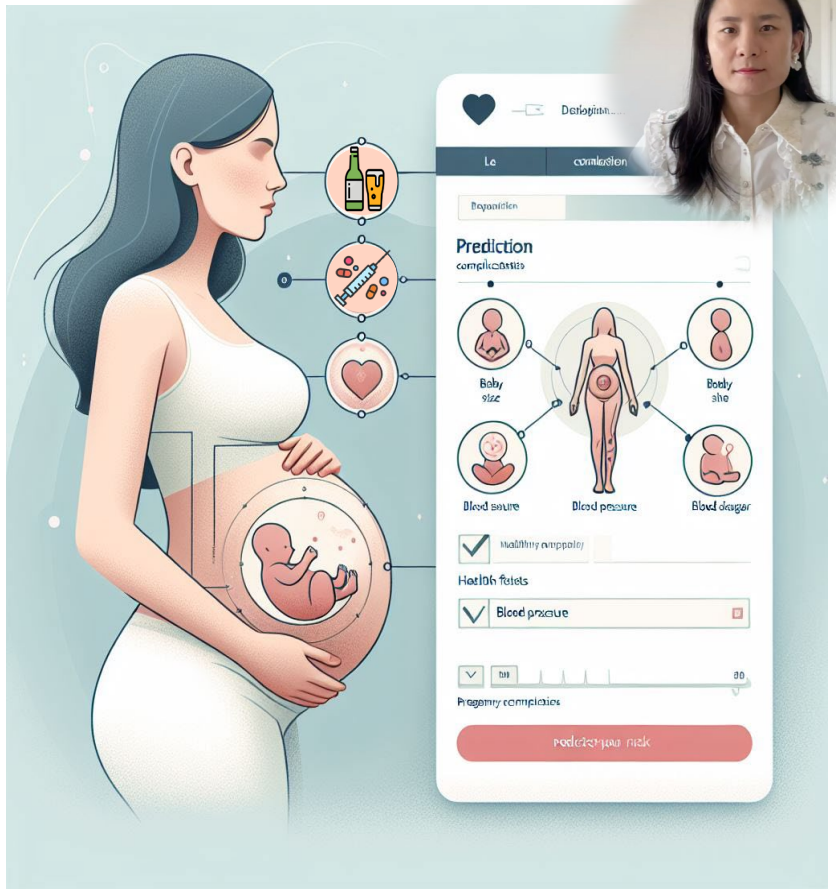
HELLO

USING MACHINE LEARNING TO PREDICT THE EFFECT OF ALCOHOL AND SUBSTANCE USE ON WOMEN AND BABIES' HEALTH



Yanan Hu

PhD Candidate at Monash University
Research Associate at University of Technology Sydney

Yanan.hu@monash.edu

Acknowledgment of Country

I wish to acknowledge the people of the Kulin Nations, on whose land we are gathered today. I pay my respects to their Elders, past and present.

Table of contents

Illustrative example: CSAI model

Define the outcome measure

1

Caesarean Section Following Induction of Labour

Collect and examine data

3

Routinely Collected Perinatal Datasets

Deal with missing data

5

- Grouped as unknown for categorical variables
- Assigned an unrealistic outlier value for continuous variables

2

Identify candidate predictors

N=18
- BMI
- Age
- Rurality
- SEIFA
- Parity

- Smoke
- Diabetes
- Birthplace
- .....

- Alcohol use
- Substance use

4

Consider sample size

180,700

6

Fit the prediction models

7 Machine Learning Classification Algorithms

Efthimiou O, Seo M, Chalkou K, Debray T, Egger M, Salanti G. Developing clinical prediction models: a step-by-step guide. *BMJ* 2024; **386**: e078276.
Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020; 368: m441

Assess the model performance

7

- Discrimination
- Calibration
- Confusion matrix
- .....

Perform a decision curve analysis [1]

9

Make the model accessible to potential users

11

8

- Performance
- Number of included predictors (n=10)
- Training time (1h)

Decide on the final model

10

Contribution of each predictor

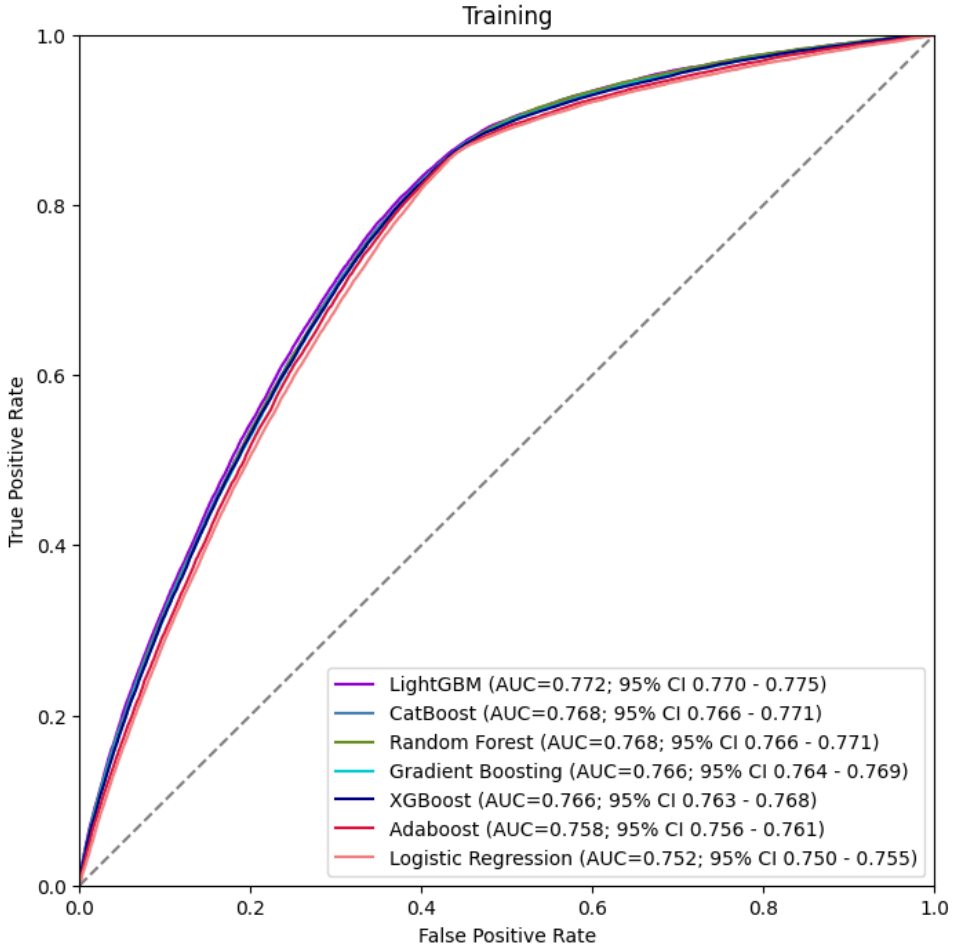
SHAP [2]

12

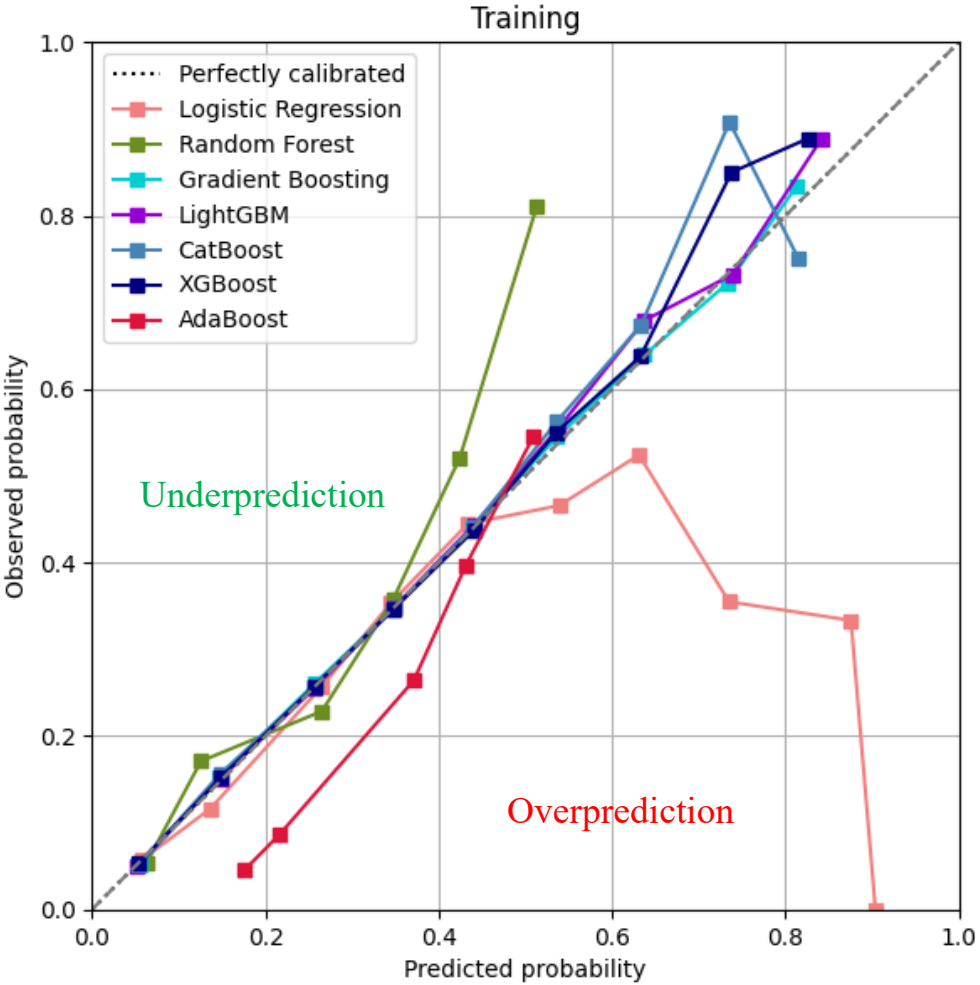
Write up and publish

1. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006; 26(6): 565-74.
2. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in neural information processing systems 2017; 30.

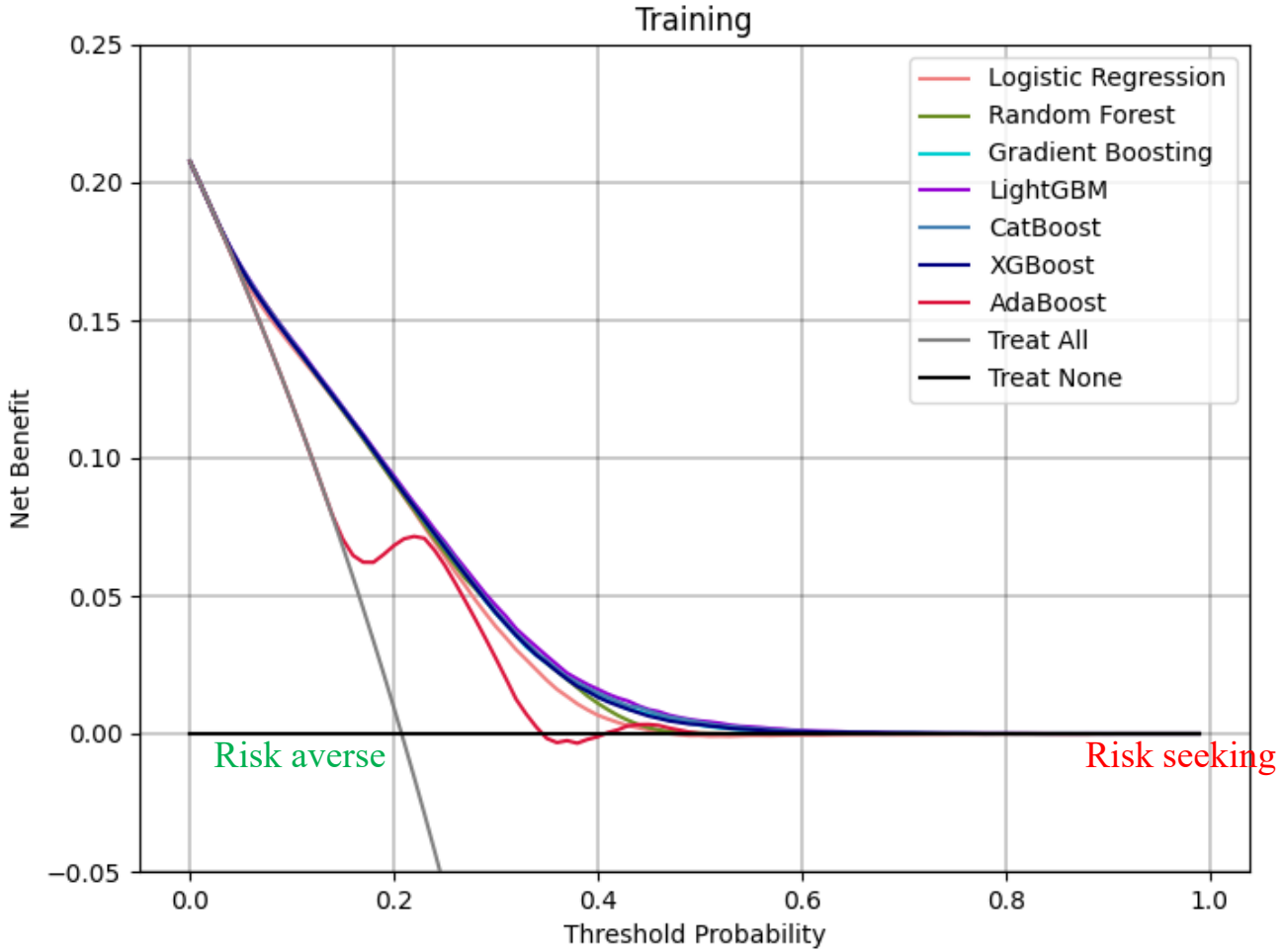
RECEIVER OPERATING CHARACTERISTIC CURVE (ROC) CURVE



CALIBRATION CURVE



DECISION CURVE ANALYSIS



Predicting Caesarean Section Following Induction of Labour (CSAI Web App)

Important Information:

This calculator is only applicable to women who consider induction of labour and meet the eligibility criteria (singleton, cephalic, term, and no prior history of caesarean section).

It must be used in addition to, and not as a substitute for, appropriate clinical judgement.

Instructions:

All fields need to be answered. Default values have been pre-filled for demonstration purposes only.

For fill-in-the-blank questions except postcode, if the information is not known or uncertain, please enter 99. For an unknown postcode, please enter 9999.

Hover over the information icon i for more details.

Maternal age (years): 31

Pre-pregnancy BMI (kg/m²): i 26

Postcode of woman's primary residence: i 3151

Planned weeks of gestation for induction of labour: i ○ 37 ○ 38 ◉ 39 ○ 40 ○ 41

Have you given birth before: i ○ Yes (with prior vaginal birth) ◉ No ○ Not known

Giving birth in a private hospital: ○ Yes ◉ No ○ Not known

Smoked *before* 20 weeks of gestation: ○ Yes ◉ No ○ Not known

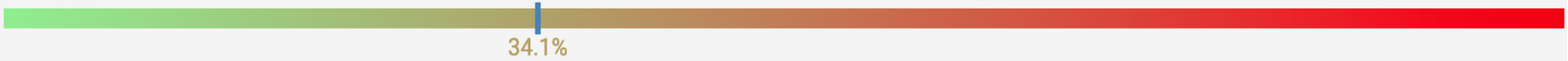
Pre-existing diabetes before current pregnancy: i ○ Yes ◉ No/Not known

Preeclampsia during current pregnancy: i ○ Yes ◉ No/Not known

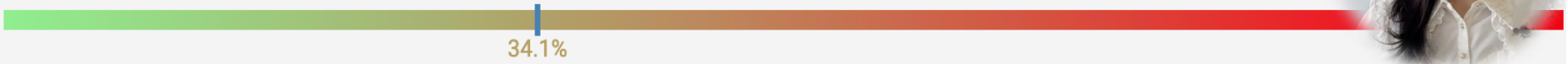
Which region the *woman* was born in: i ◉ Australia (includes External Territories) ○ Oceania and Antarctica (excludes Australia) ○ Europe ○ Africa ○ Asia ○ America ○ Not known

Predict Reset

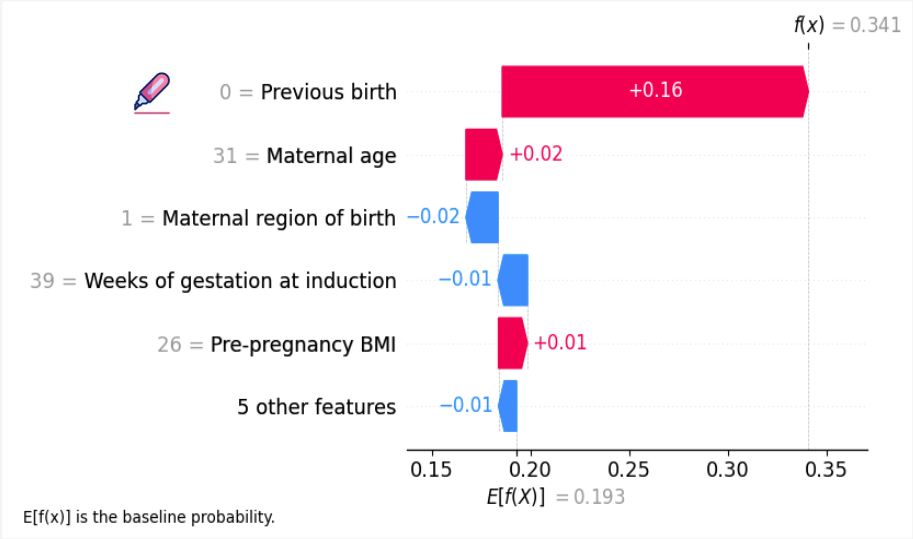
Predicted probability of caesarean section following induction of labour:



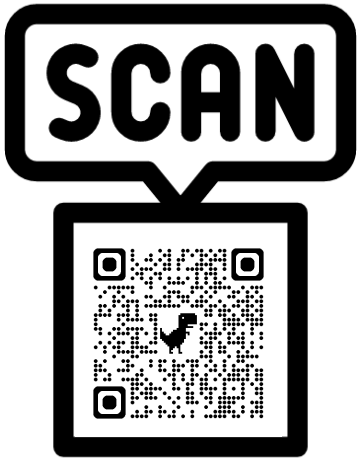
Predicted probability of caesarean section following induction of labour:



Plotting the contribution of each predictor, ranked by importance:



E[f(x)] is the baseline probability.



CSAI WEB APP

Thanks!

Any questions?

Please keep all content shared today confidential, as the paper has not yet been published.