

PREDICTING PRETERM BIRTHS USING U.S. NATIONAL BIRTH DATA: A DEEP LEARNING APPROACH

Zhu Shuying, Dellinger Andrew, Chan Karen, Lam Wendy, Pang Herbert

BACKGROUND

Preterm Births

- Gestational age at birth < 37
- Occurs in 5% to 18% of births worldwide
- Cause of neonatal morbidity and mortality globally
 - Result in long-term health problems
 - developmental delay in childhood
 - cardiometabolic disorders in the adulthood of survivors
- Predicting preterm births assists:
 - risk assessments for timely referrals
 - enable intervention strategies for women in their early pregnancy.

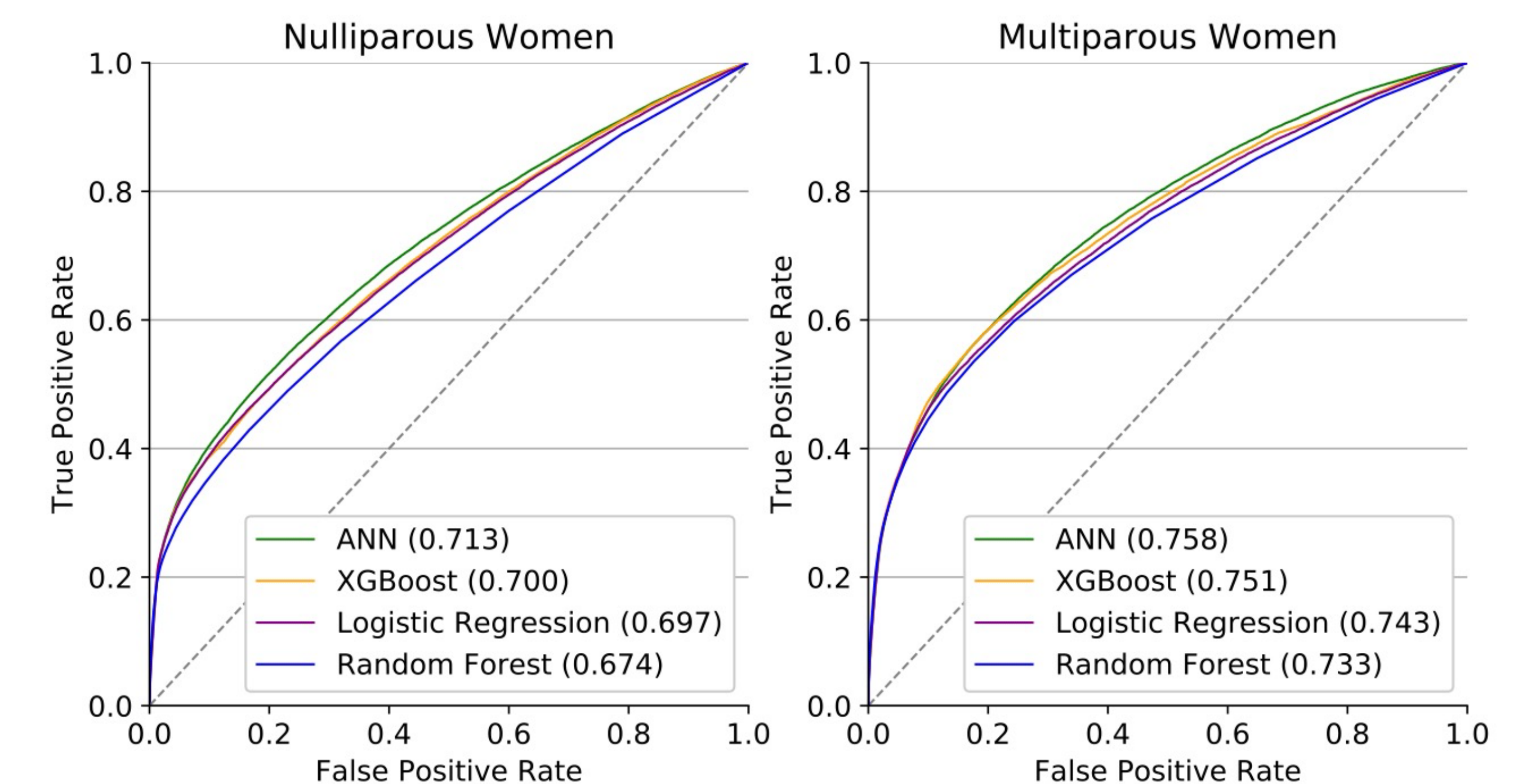
Gaps of Existing Research

- Small data set
- Use logistic regression which assumes the underlying relationship is linear.
- The models demonstrated low sensitivity (4.2%-44%)

METHODS

- Artificial neural networks with modern deep learning techniques were developed using national birth data (n = 17,378,139) of women who began their first prenatal visit no later than week 24 of pregnancy in the United States between 2014 and 2018 to predict preterm births.
- For the prediction, 26 and 34 routinely collected variables that are known before the 24 weeks of pregnancy were used for nulliparous women and multiparous women, respectively.
- Training, validation, and test data sets were used to build models and assess performance. Predictive performance was assessed by AUC.
- The proposed models were compared with logistic regression and other machine learning methods. Additional analysis was performed to identify important variables.

RESULTS



- Our ANN model outperforms XGBoost, LR and RF for both nulliparous women and multiparous women.

- Our model reaches the sensitivity of 68%.



**HKU
Med**

LKS Faculty of Medicine
The University of Hong Kong
香港大學李嘉誠醫學院