Development and validation of ensemble machine-learning based web-embedded decision supporting tool for prostate biopsy

Jungyo Suh[1], *Sangjun Yoo*[2], *Juhyun Park*[2], *Sung Yong Cho*[1], *Min Chul Cho*[2], *Hwancheol Son*[2], *Hyeon Jeong*[2]

1. Department of Urology, Seoul National University hospital college of medicine
2. Department of Urology, SMG-SNU Boramae Medical Center

American Urological Association

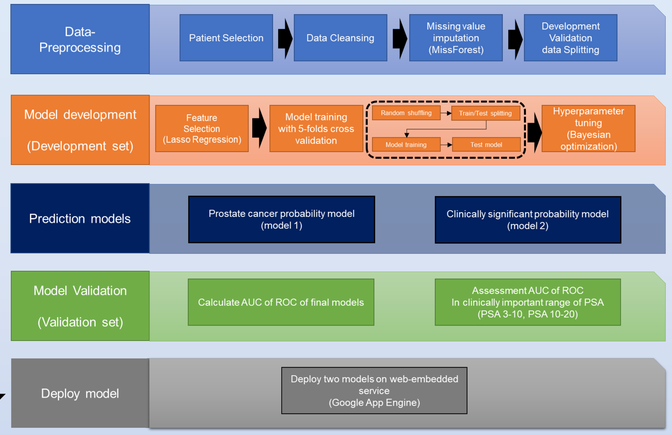
AUA-2020 MAY 15-18 washington, dc

INTRODUCTION

- There are several multivariable prediction models for diagnosing prostate cancer (PC) and clinically significant prostate cancer (csPC) none of them exhibit outstanding performance than PSA alone.
- **Complex machine learning model makes better prediction with reflect complexity of real-world correlation of features.**
- However, **complex machine learning model has problem in black-box phenomena for clinical application.**
- In this study, we developed and validated the **explainable artificial intelligence (XAI) machine model (XGBoost)** for calculating the probability of PC and csPC and deployed it using a web-embedded structure for clinician's decision support prior to prostate biopsy.

MATERIALS & METHODS

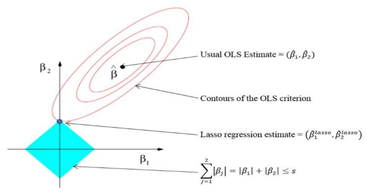
Overall workflow



Patients information

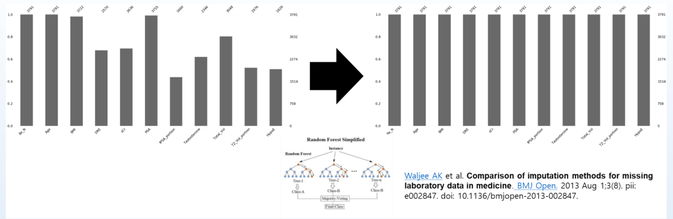
- **3832 prostate biopsy** patients' database
- Since March 2009 to October 2019
- **Included parameters on DB:** Age, Biopsy history, BMI, PSA level, free PSA level, Testosterone level, serum creatinine level, DRE abnormality, total prostate volume, transitional zone prostate volume, hypoechoic lesion on ultrasound

Feature selection

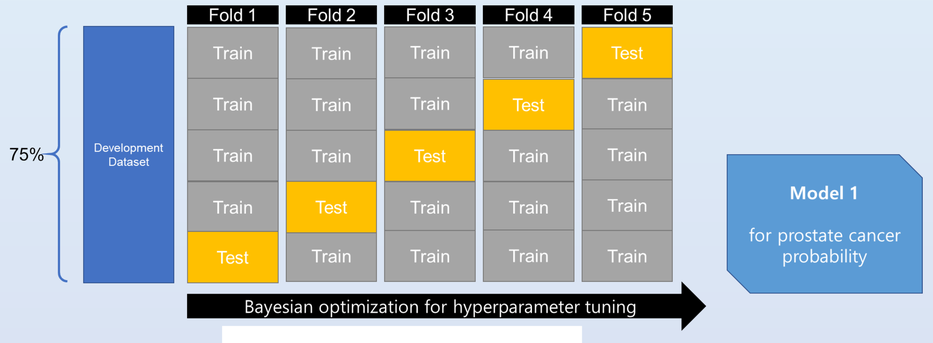


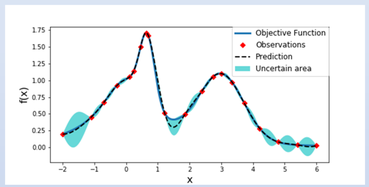
- Lesser parameter makes better application
- We reduced parameters using LASSO (least absolute shrinkage and selection operator) regression for each models

Missing Value Imputation

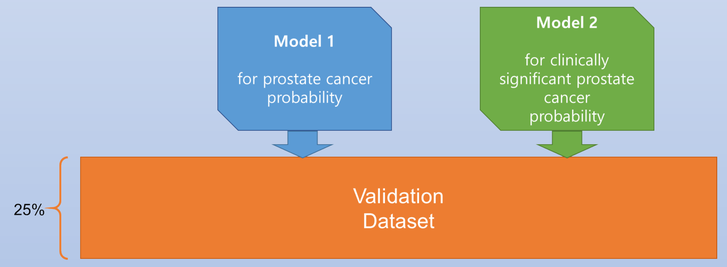


Model Development





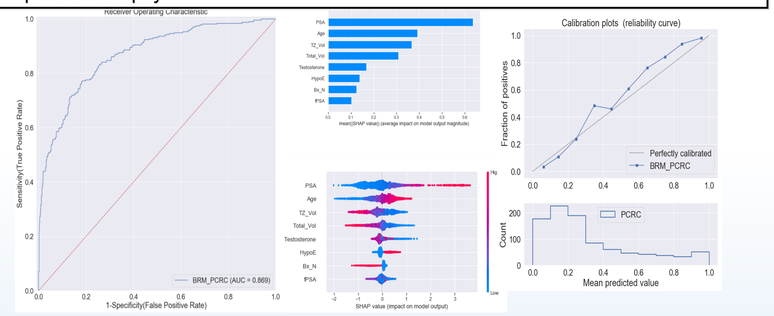
Model Validation



RESULTS

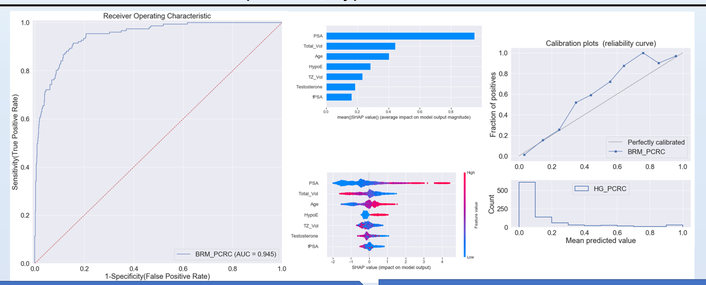
Model 1 Performance

- Parameters (8) : Age, PSA, free PSA, Testosterone, Total volume prostate, transitional zone volume prostate, hypoechoic lesion, Previous number of prostate biopsy

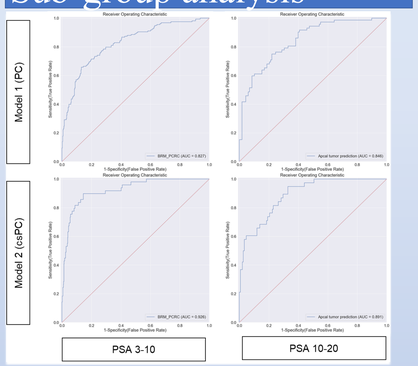


Model 2 Performance

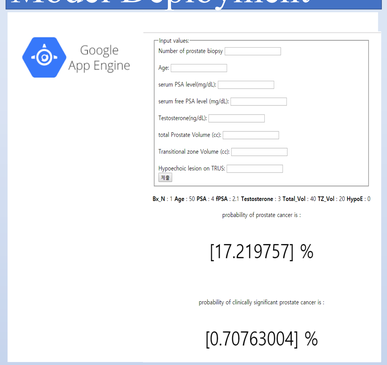
- Parameters (7) : Age, PSA, free PSA, Testosterone, Total volume prostate, transitional zone volume prostate, hypoechoic lesion



Sub-group analysis



Model Deployment



Conclusion

- We successfully developed and validated a decision-supporting tool using XAI for calculating the probability of PC and csPC prior to prostate biopsy.
- You can access freely with URL: https://boramae-pcrc.appspot.com