Machine Learning Stacking Ensemble Model for Predicting Heart Attacks

The Eighth International Conference on Big Data, Small Data, Linked Data and Open Data

2022 - Barcelona, Spain

Muath Obaidat, Alex Alexandrou, and Samantha Sanacore
City University of New York

Presenter: Muath Obaidat
Associate Professor at the City University of New York
Contact email: muobaidat@ccny.cuny.edu

CUNY THE CITY UNIVERSITY OF NEW YORK

IARIA

Muath A. Obaidat

- Received his Ph.D. from the City University of New York in Electrical Engineering. Prior to that, he earned his Master's degree in Computer Engineering from New Jersey Institute of Technology. In addition, he holds a degree in Computer Science from Monmouth University in New Jersey.
- Currently he is an Associate Professor in Cybersecurity and Digital Forensics at the City University of New York
- His research interests lie in the area of cybersecurity, digital forensics, ubiquitous Internet of Things (IoT) security and privacy. His recent research includes Machine learning and security.

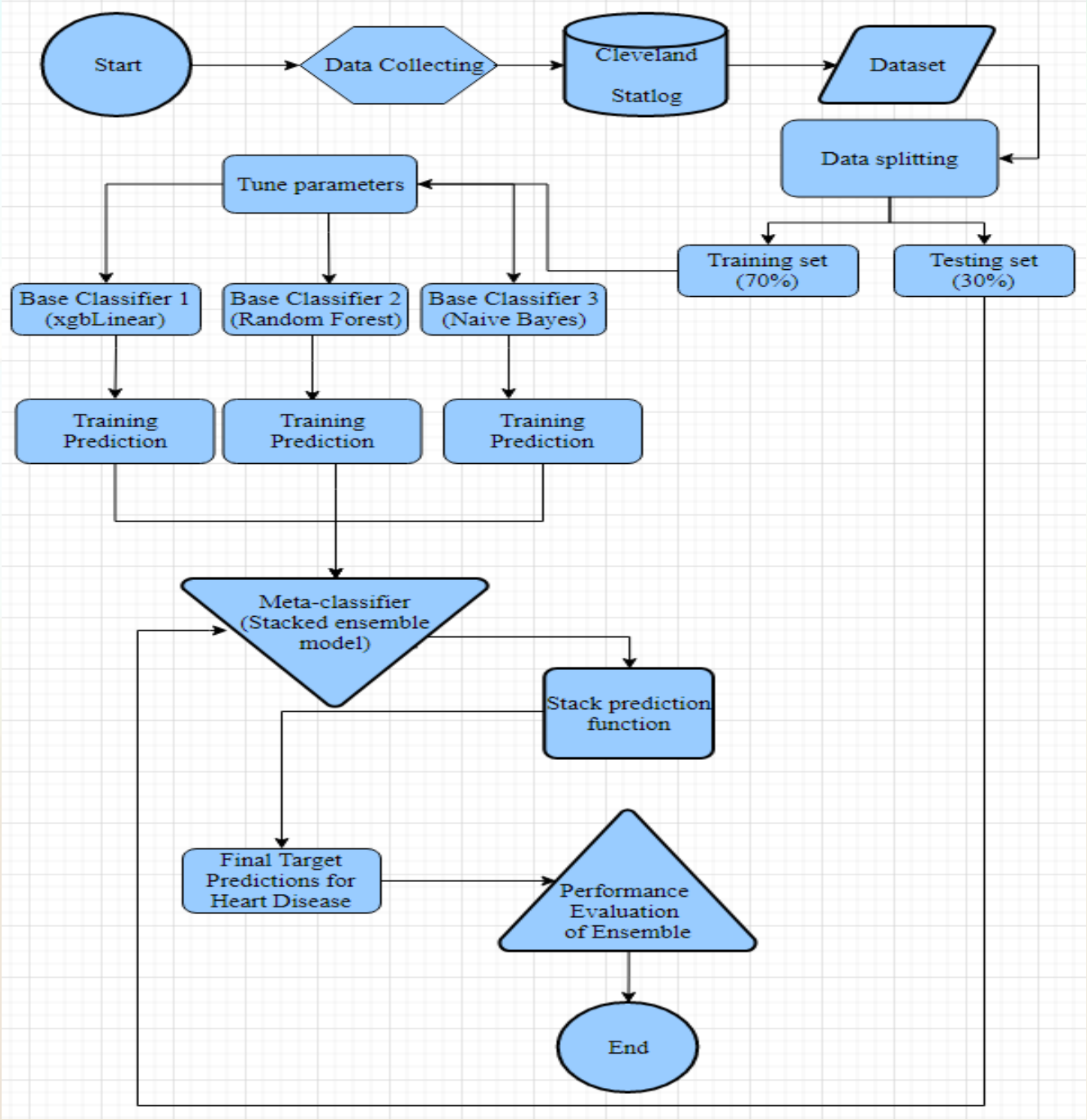
CUNY THE CITY UNIVERSITY OF NEW YORK

IARIA

Goal and Contribution

- The goal is to develop a Machine Learning model to reach a very high level of accuracy for predicting myocardial infarction, known as heart attack.
- The contribution of this study is to develop a Machine learning model for predicting heart attack.
- Also, studying the performance of the proposed scheme against the benchmarks in the literature.

Proposed Scheme



THE CITY UNIVERSITY OF NEW YORK

IARIA

Stacked Ensemble Model

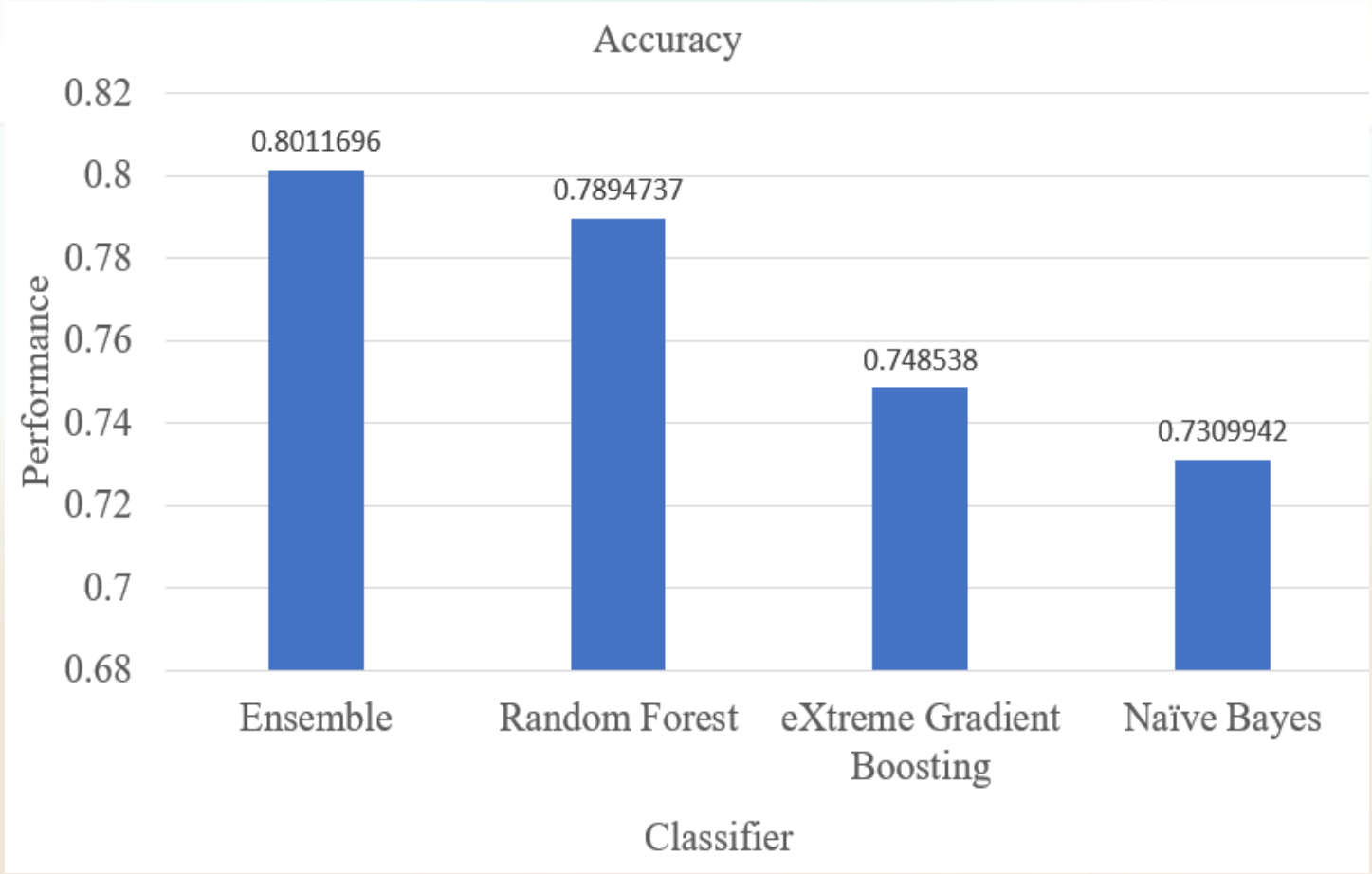
- Naive Bayes classifier
  - ✓ $P(A|B) = \frac{P(B|A)*P(A)}{P(B)}$
- $P(X|y) = P(x_1|y) * P(x_2|y) * \cdots * P(x_n|y)$
- $y = \text{argmax}_y[P(y) * \prod_{i=1}^{n} P(x_i|y)]$
- a new data set that contains $\{x_i^{new}, y_i\}$
- $X_i^{new} = \{h_1(x_i), h_2(x_i), \ldots, h_T(x_i)\}$

- Fianl $H(x) = h^{new}(h_1(x), h_2(x), \ldots, h_T(x))$

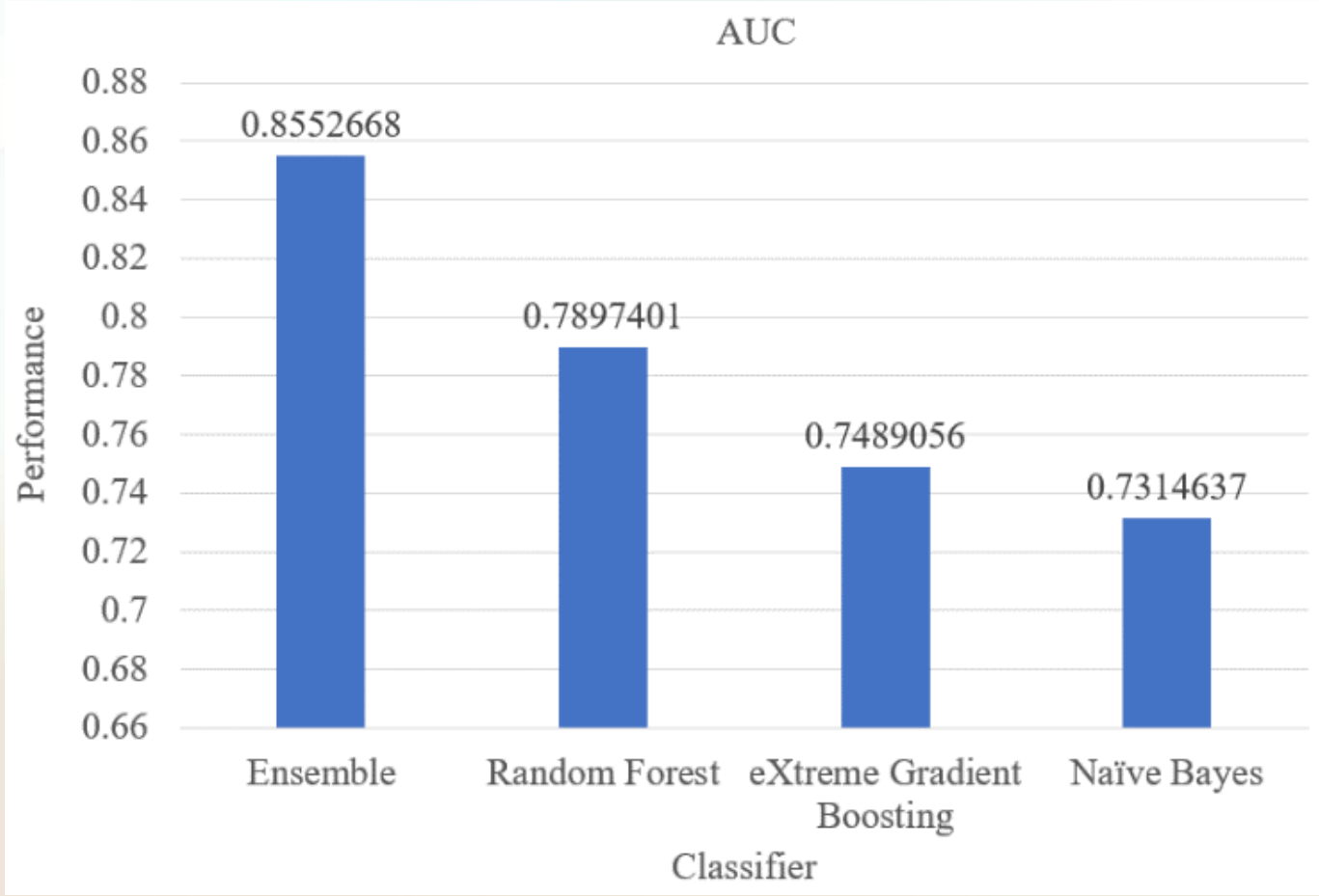
Performance Evaluation

- The metrics used to evaluate the performance of the proposed scheme are the following:
- ✓ Accuracy
- ✓ Area Under the Curve (AUC)
- ✓ Specificity
- ✓ Precision
- ✓ Sensitivity

Accuracy



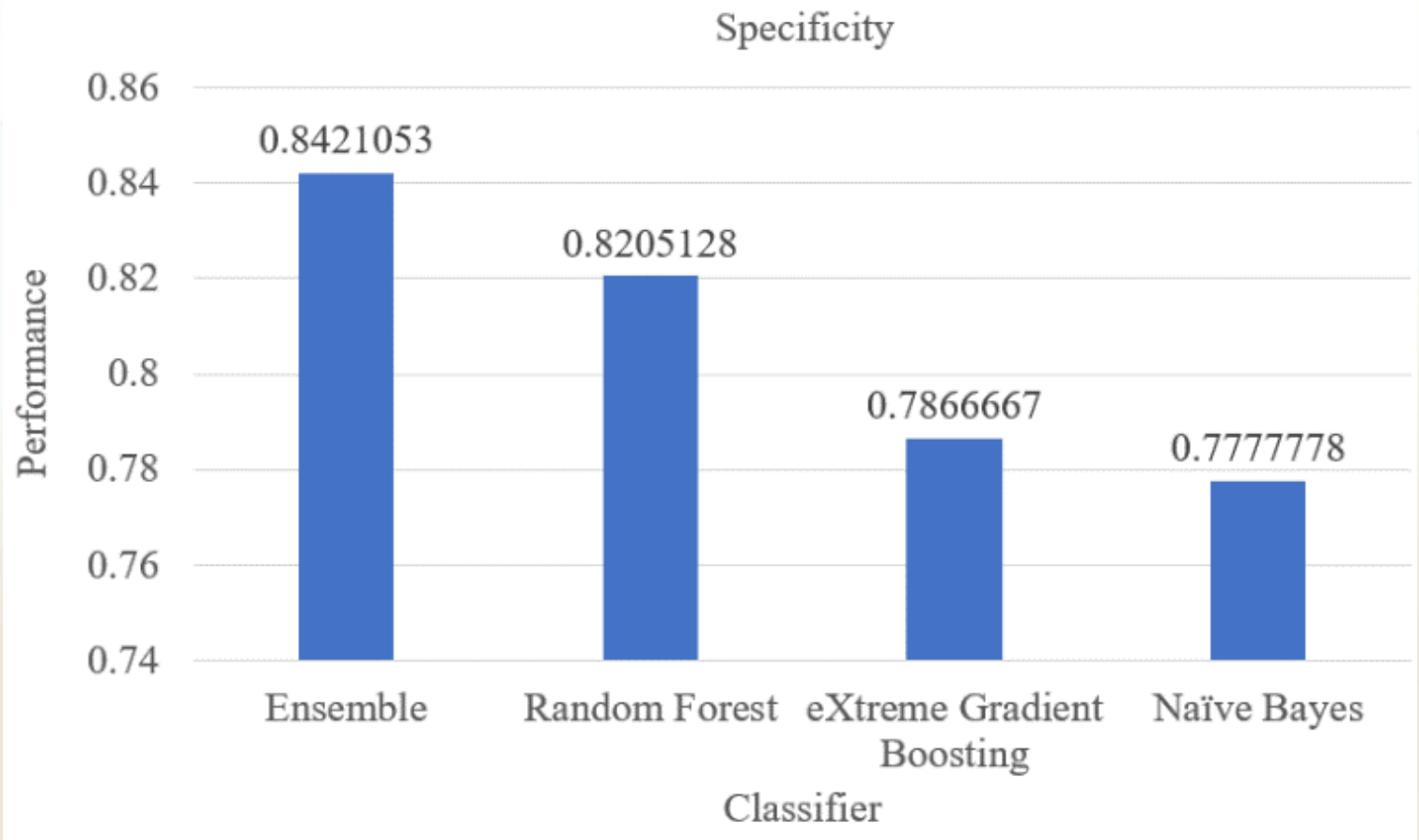
Area Under the Curve (AUC)



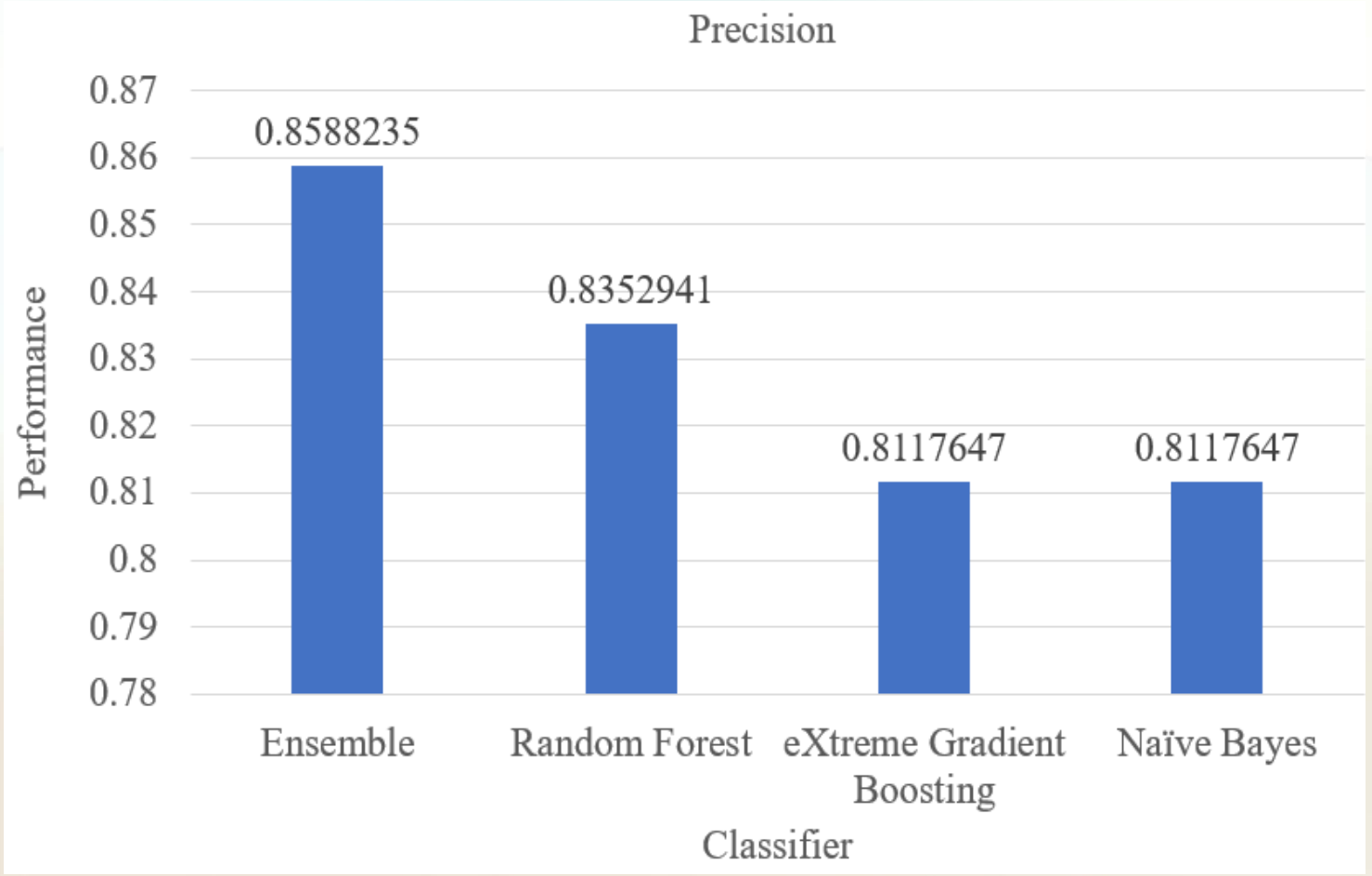
CUNY THE CITY UNIVERSITY OF NEW YORK

IARIA

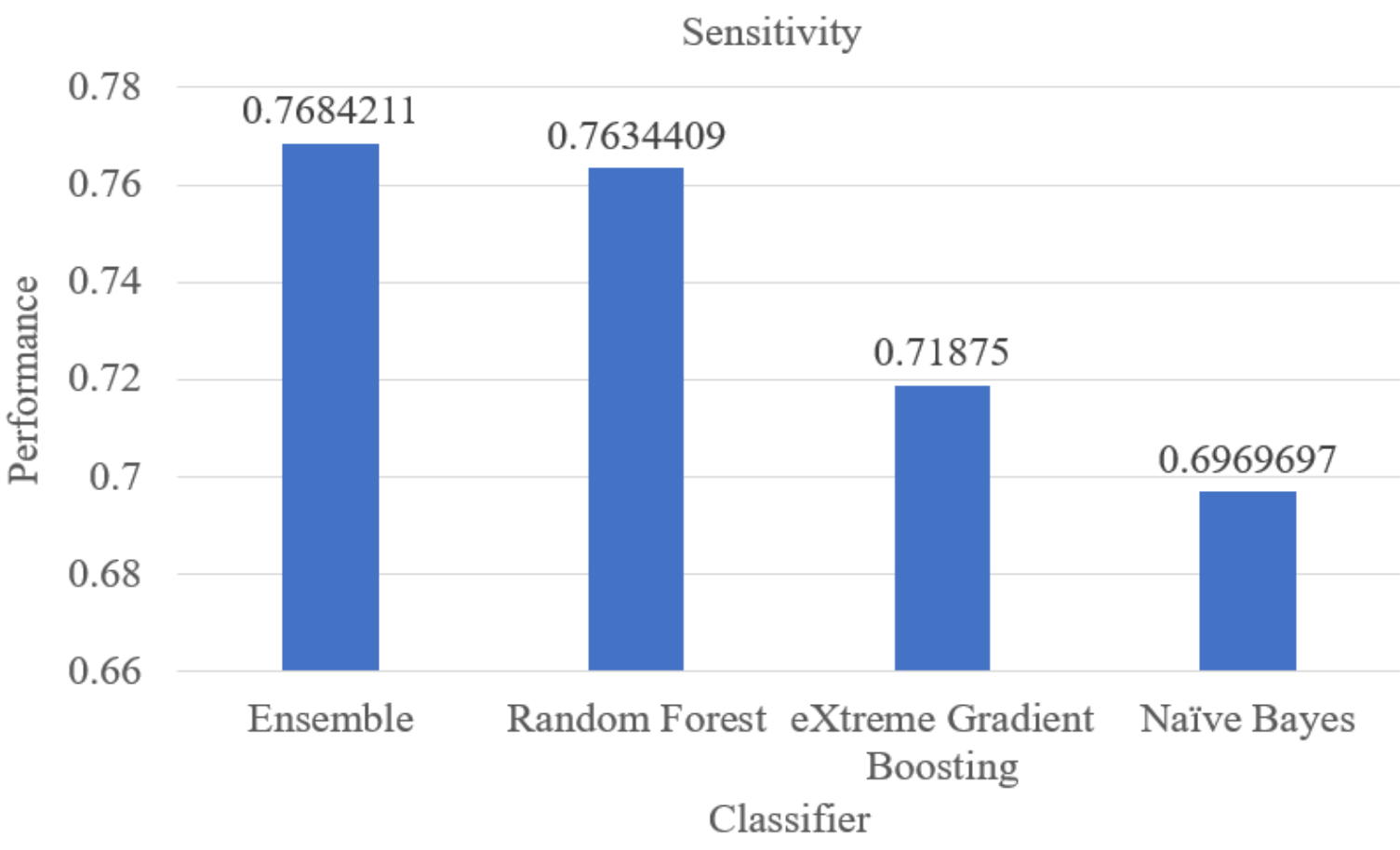
Specificity



Precision



Sensitivity



Conclusion

- This work proposed a stacked ensemble model to improve the level of accuracy for the prediction of heart attacks.
- The stacked ensemble model achieved the best performance in terms of accuracy, area under the curve, specificity, precision, and sensitivity metrics.
- The proposed model has the highest accuracy performance, which was 1.16% higher than the next best performance (i.e. Random Forest)
- The proposed model can help reduce the number of deaths caused by heart attacks worldwide

Thank you!

Q & A

CUNY THE CITY UNIVERSITY OF NEW YORK

IARIA