

FADE: FAir Double Ensemble Learning for Observable and Counterfactual Outcomes

Alan Mishler
alan.mishler@jpmorgan.com
J.P. Morgan AI Research
New York, USA

Edward H. Kennedy
edward@stat.cmu.edu
Carnegie Mellon University
Pittsburgh, USA

ABSTRACT

Methods for building fair predictors often involve tradeoffs between fairness and accuracy and between different fairness criteria. Recent work seeks to characterize these tradeoffs in specific problem settings, but these methods often do not accommodate users who wish to improve the fairness of an existing benchmark model without sacrificing accuracy, or vice versa. These results are also typically restricted to observable accuracy and fairness criteria. We develop a flexible framework for fair ensemble learning that allows users to efficiently explore the fairness-accuracy space or to improve the fairness or accuracy of a benchmark model. Our framework can simultaneously target multiple observable or counterfactual fairness criteria, and it enables users to combine a large number of previously trained and newly trained predictors. We provide theoretical guarantees that our estimators converge at fast rates. We apply our method on both simulated and real data, with respect to both observable and counterfactual accuracy and fairness criteria. We show that, surprisingly, multiple unfairness measures can sometimes be minimized simultaneously with little impact on accuracy, relative to unconstrained predictors or existing benchmark models.

CCS CONCEPTS

• **Computing methodologies** → **Ensemble methods.**

KEYWORDS

fairness, counterfactual, ensemble learning, semiparametric

ACM Reference Format:

Alan Mishler and Edward H. Kennedy. 2022. FADE: FAir Double Ensemble Learning for Observable and Counterfactual Outcomes. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAcT '22)*, June 21–24, 2022, Seoul, Republic of Korea. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3531146.3533167>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ACM FAcT '22, June 21–24, 2022, Seoul, South Korea

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9352-2/22/06.

<https://doi.org/10.1145/3531146.3533167>