

JOSE M. ALVAREZ

University of Pisa, Department of Computer Science, 56127, Italy
351 522 9588 ◊ jose.alvarez@di.unipi.it ◊ cc-jalvarez.github.io

RESEARCH STATEMENT

My work focuses on using causality-based modeling frameworks for formalizing, detecting, and mitigating bias in automated decision-making processes. I am particularly interested in finding new methods for enhancing existing causality models to operationalize the social meanings we attribute to protected attributes—such as gender and race—that are often not captured by data alone but are yet essential in framing what we mean by (un)fair decision outcomes within a specific context.

Key words: machine learning, fairness, structural causal models, discrimination analysis

EDUCATION

Scuola Normale Superiore Ph.D. Data Science, jointly offered with the University of Pisa <i>Advisor:</i> Salvatore Ruggieri	Nov 2020 - Present
Toulouse School of Economics M.Sc., Econometrics and Empirical Economics	Sep 2015 - May 2017
University of Florida B.A., Economics and History, <i>summa cum laude</i>	Sep 2011 - May 2015

EXPERIENCE

University of Pisa <i>Early-Stage Researcher with NoBIAS ITN</i>	Nov 2020 - Present <i>Pisa, IT</i>
<ul style="list-style-type: none">• NoBIAS: Artificial Intelligence without Bias is part of the Marie Skłodowska-Curie Innovative Training Network (ITN) and funded by the European Union's Horizon 2020• The Marie Skłodowska-Curie ESR position is through the University of Pisa's Knowledge Discovery and Data Mining Laboratory (KDD Lab)	
KU Leuven <i>Visiting Fellow</i>	Sep 2022 - Nov 2022 <i>Leuven, BE</i>
<ul style="list-style-type: none">• Visited under Prof. Bettina Berendt, Computer Science Department, the The Declarative Languages and Artificial Intelligence (DTAI) Section	
Generali <i>Visiting Research Scientist</i>	Sep 2021 - Jan 2022 <i>Milan, IT</i>
<ul style="list-style-type: none">• Worked with Human Resources in understanding and operationalizing the current human-driven hiring process using fair ranking modeling techniques	
Deloitte Risk Advisory <i>Senior Consultant - Operational & Financial Risk Modeler</i>	May 2017 - Oct 2020 <i>Brussels, BE</i>
<ul style="list-style-type: none">• Designed, developed, and deployed statistical learning models for clients in the energy & utilities and financial sectors• Developed an end-to-end time series forecasting tool at the Deloitte aiStudio in Frankfurt, Germany, as the chosen Belgium representative• Conducted internal and external trainings in R programming, Git, Advanced Loss Given Default modeling (Basel II), and Survival Analysis modeling	

- Provided insights via statistical analysis to policy makers on a range of topics, such as studying the volunteering determinants in the UK and the variation in electoral costs across local authorities

PUBLICATIONS

Salvatore Ruggieri, **Jose M. Alvarez**, Andrea Pugnanna, Laura State, Franco Turini. “Can We Trust Fair-AI?” *AAAI Conference on Artificial Intelligence*, 2023.

Matilde Lazzari, **Jose M. Alvarez**, Salvatore Ruggieri. “Predicting and explaining employee turnover intention.” *International Journal of Data Science and Analytics*, 2022, [paper].

WORKING PAPERS

Jose M. Alvarez, Salvatore Ruggieri. “Counterfactual Situation Testing: Fairness given the Difference.” *In submission*, [paper].

Jose M. Alvarez, Mayra Russo. “Perception as a Fairness Parameter” [extended abstract].

Carlos Mougán, **Jose M. Alvarez**, Gourab K. Patro, Salvatore Ruggieri, Steffen Staab. “Fairness implications of encoding protected categorical attributes.” [paper].

INVITED TALKS, POSTERS, AND WORKSHOPS

2022

- ACM Conference on Fairness, Accountability, and Transparency (FAccT): *doctoral colloquium*;
- European Workshop on Algorithmic Fairness (EWAf): *invited talk*;
- Comete Workshop on Ethical AI: *invited talk*;
- NeurIPS Workshop on Algorithmic Fairness through the Lens of Causality and Privacy: *poster*.

PROFESSIONAL SERVICES

Reviewer for *NeurIPS 2022 Workshop on Algorithmic Fairness through the Lens of Causality and Privacy (AFCP 2022)*.

Sub-reviewer for *AI*IA 2021, ECML-PKDD 2021, IJCAI 2021, UAI 2021, AISTATS 2021, AISTATS 2022, WWW 2022, Data Mining and Knowledge Discovery Journal, AISTATS 2023, TheWebConf 2023*.

FELLOWSHIPS, GRANTS, AND AWARDS

Marie Skłodowska-Curie Fellowship	2020 - 2023
Phi Beta Kappa	2015
University of Florida's Honors Program	2015
Davis United Word College Scholar	2011 - 2015

SKILLS

Programming	Python, R, Scala
Other	Git, L ^A T _E X, Stata, SQL
Languages	Spanish, English, Italian, French