



Topic

Innovating AI Evaluation: Beyond Accuracy and Precision.

Description

As the landscape of artificial intelligence continues to evolve, the need for comprehensive and nuanced evaluation methods increases as well. Traditional metrics such as accuracy and precision, while important, are insufficient for fully capturing the complexities and impacts of AI systems.

The SAIL Spring School aims to address this gap by introducing participants to a diverse array of evaluation strategies, such as user evaluations, ethical and societal impacts, evaluating outcomes that are co-constructed between user and AI, mathematical guarantees, interpretability and transparency assessments, context-specific metrics, etc.

More information and registration under: www.sail.nrw/springschool/

When & Where

When: March 26-28, 2025

Where: CITEC lecture hall, Bielefeld University, Germany

Fair Reinforcement Learning



Mykola Pechenizkiy
Eindhoven University of Technology



Yingqian Zhang
Eindhoven University of Technology

In this tutorial we look at algorithmic fairness in the context of sequential decision-making across various high-impact real-world domains. We will start with providing a comparative analysis of the fairness conceptualizations in the (relatively mature) field of fair-ML focusing on supervised learning scenarios with the (re-)emerging field of fair-RL traditionally focusing of reinforcement learning for resource allocation. Next, we will review different notions of fairness in fair-RL and discuss integrating fairness into fair-RL as a multi-objective optimization problem, and examine the trade-off between fairness and utility with several existing studies on various applications. We will conclude with a discussion of future challenges in fair-RL and their prospective solutions, exploring interdisciplinary insights from successful fair-ML research and applying them to the realm of fair-RL.

Funded by

