The Alan Turing Institute
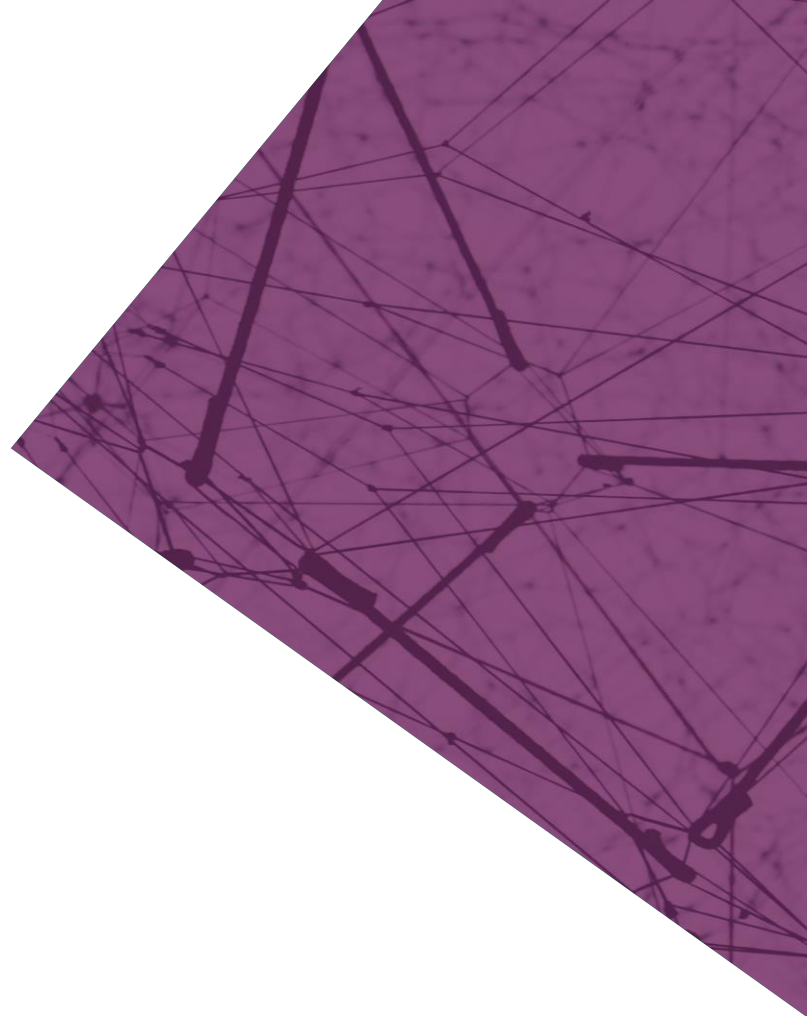

Milestone 3

Sources, Forms, and Quantification of Bias and Discrimination in Supervised Learning

Roseline Polle

Postgraduate, UCL

roseline.polle.19@ucl.ac.uk

Course Overview

Part I: Ethical, Legal, and Governance

Milestone 1. Notions of Fairness

Milestone 2. Legal and Governance aspects of AI Fairnes

Part II: Technical dimensions in Supervised ML

Milestone 3. Sources, Forms, and Quantification

Milestone 4. Mitigation

Milestone 5. Trade-offs and Interactions with other verticals in Trustworthy AI

Milestone Overview

- **Part 1: Theory** → Slides + Supporting notebook
    1. Real-life examples of bias in algorithms
    2. Type and sources of Bias
    3. Quantification of Bias: Fairness definitions and metrics
    4. Choosing a metric

- **Part 2: Practice** → Notebooks
    1. Explore the data
    2. Quantify bias manually
    3. Quantify bias using aif360
    4. Getting a confidence interval