Algorithmic Bias and Fairness Risk Workshop Training Course

Professional Development Training Course Outline

Category	Risk Management	**Duration**	5 Days
Base Fee	$1,500 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

The rapid global adoption of Artificial Intelligence (AI) and Machine Learning (ML) systems has made them central to critical decision-making across sectors, from finance and hiring to healthcare and criminal justice. However, these systems inherently carry the risk of perpetuating and even amplifying existing societal prejudices, a phenomenon known as algorithmic bias. This can lead to non-compliant, unethical, and harmful outcomes, causing significant reputational risk and eroding public trust in technology. Organizations must move beyond mere compliance with emerging regulations like the EU AI Act and implement a proactive, systemic approach to Responsible AI (RAI) and AI Governance.

Algorithmic Bias and Fairness Risk Workshop Training Course is designed to equip participants with the technical, legal, and ethical frameworks necessary to effectively identify, quantify, and mitigate algorithmic discrimination and fairness risk throughout the entire AI/ML lifecycle. We will explore the various sources of bias including data bias, systemic bias, and feedback loops and provide practical, hands-on strategies for implementing fairness-aware machine learning and ensuring algorithmic transparency and accountability. The goal is to build a culture of Ethical AI that fosters innovation while safeguarding digital equity and adhering to the highest standards of data ethics.

Programme Curriculum

Algorithmic Bias and Fairness Risk Workshop Training Course

Introduction

The rapid global adoption of Artificial Intelligence (AI) and Machine Learning (ML) systems has made them central to critical decision-making across sectors, from finance and hiring to healthcare and criminal justice. However, these systems inherently carry the risk of perpetuating and even amplifying existing societal prejudices, a phenomenon known as algorithmic bias. This can lead to non-compliant, unethical, and harmful

outcomes, causing significant reputational risk and eroding public trust in technology. Organizations must move beyond mere compliance with emerging regulations like the EU AI Act and implement a proactive, systemic approach to Responsible AI (RAI) and AI Governance.

Algorithmic Bias and Fairness Risk Workshop Training Course is designed to equip participants with the technical, legal, and ethical frameworks necessary to effectively identify, quantify, and mitigate algorithmic discrimination and fairness risk throughout the entire AI/ML lifecycle. We will explore the various sources of bias including data bias, systemic bias, and feedback loops and provide practical, hands-on strategies for implementing fairness-aware machine learning and ensuring algorithmic transparency and accountability. The goal is to build a culture of Ethical AI that fosters innovation while safeguarding digital equity and adhering to the highest standards of data ethics.

Course Duration

5 days

Course Objectives

1. Establish a clear understanding of algorithmic bias, discrimination, and fairness
2. Pinpoint the primary sources of bias in the AI/ML lifecycle, including data collection bias and labeling bias.
3. Master key fairness metrics and their trade-offs.
4. Implement a structured methodology for AI Impact Assessment (AIIA) and fairness risk scoring.
5. Learn data debiasing techniques such as reweighing, sampling, and data augmentation.
6. Utilize fairness-aware learning and constrained optimization in model training.
7. Implement prediction-adjustment techniques like score calibration and thresholding.
8. Apply Explainable AI (XAI) methods to interpret model decisions and audit for bias.
9. Analyze the requirements for AI Governance and risk management under emerging frameworks.
10. Design an effective continuous monitoring and bias audit framework for deployed AI systems.
11. Identify and mitigate new forms of bias and harm specific to Large Language Models (LLMs) and Generative AI.
12. Understand the crucial role of diverse teams and stakeholder engagement in bias detection and mitigation.
13. Create a practical, context-specific Responsible AI roadmap for their organization's specific use cases.

Target Audience

1. Data Scientists & Machine Learning Engineers
2. AI Product Managers
3. AI/ML/Data Ethics Governance Committees
4. Compliance, Risk, and Legal Officers
5. Chief Data/Analytics Officers
6. UX/UI Designers for AI Applications
7. Data Quality and Data Audit Specialists
8. Senior Business Leaders/Decision Makers

Course Modules

Module 1: Foundational Concepts & Ethical Imperative

- Defining Algorithmic Bias

- The spectrum of Fairness Definitions
- Mapping bias to real-world harm
- The Business Case for Fairness
- Case Study: The COMPAS Recidivism Prediction tool and the ProPublica investigation into racial disparity in error rates.

Module 2: Sources and Types of Bias in the AI Lifecycle

- Historical, Selection, Sampling, Measurement, and Labeling Bias.
- Feature selection/engineering and objective function design.
- Context-shift and human-in-the-loop misuse.
- How model outcomes reinforce initial biases.
- Case Study: The Amazon AI Recruiting Tool.

Module 3: Fairness Metrics and Quantitative Assessment

- Introduction to Protected Attributes and Sensitive Features.
- Group Fairness Metrics.
- Understanding the Incompatibility of Fairness Definitions
- Tools for quantitative fairness evaluation
- Case Study: Racial bias in U.S. Healthcare Algorithms and the failure of simple accuracy metrics.

Module 4: Pre-processing Bias Mitigation Techniques

- Techniques for Data Reweighing and Oversampling minority groups.
- Feature modification and representation learning approaches.
- Identifying and addressing proxy variables and spurious correlations.
- Data quality assurance and establishing Fairness Audits on raw data.
- Case Study: Using data rebalancing and synthetic data generation to address under-representation in a loan application dataset.

Module 5: In-processing and Post-Processing Mitigation

- Modifying the learning objective and adversarial debiasing.
- Calibrating model scores and adjusting classification thresholds for disparate impact.
- Trade-offs between Fairness and Accuracy
- Implementing Human-Centric Review in high-stakes decisions.
- Case Study: Applying threshold adjustment to a credit scoring model to achieve Equal Opportunity for different demographic groups.

Module 6: Algorithmic Transparency and Explainability (XAI)

- The need for Algorithmic Transparency and the Right to Explanation.
- Using XAI Techniques to diagnose bias in black-box models.
- Creating Model Cards and Data Sheets for Datasets for documentation and accountability.
- Interpreting feature importance to reveal discriminatory features.
- Case Study: Explaining a biased criminal justice risk score using SHAP values to reveal reliance on protected attributes

Module 7: AI Governance and Regulatory Risk Management

- Introduction to Responsible AI Frameworks and AI Ethics Principles.

- Overview of global AI regulations.
- Developing an AI Impact Assessment and Risk Scoring Matrix.
- Establishing Accountability Frameworks and clear roles for AI Governance.
- Case Study: Navigating regulatory requirements for a high-risk AI system

Module 8: Emerging Challenges and Future of Fairness

- Bias in Large Language Models.
- Techniques for LLM Alignment and bias mitigation.
- Addressing Intersectional Fairness
- The role of Socio-Technical Context in defining and measuring fairness.
- Case Study: Analyzing bias in a Generative AI Image Model regarding occupational stereotypes

Training Methodology

This course employs a participatory and hands-on approach to ensure practical learning, including:

- Interactive lectures and presentations.
- Group discussions and brainstorming sessions.
- Hands-on exercises using real-world datasets.
- Role-playing and scenario-based simulations.
- Analysis of case studies to bridge theory and practice.
- Peer-to-peer learning and networking.
- Expert-led Q&A sessions.
- Continuous feedback and personalized guidance.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call +254769199797

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	27 Nov 2026	Online	$1,000
23 Nov 2026	27 Nov 2026	Physical	$1,500
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$3,000
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$4,500
23 Nov 2026	27 Nov 2026	Physical	$6,000
23 Nov 2026	27 Nov 2026	Physical	$6,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com