Training Course on Fairness and Accountability in AI Systems

Professional Development Training Course Outline

Category	Data Science	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

The pervasive integration of Artificial Intelligence (AI) across diverse sectors has heralded unprecedented advancements, yet it simultaneously presents profound ethical challenges concerning **fairness, bias, and accountability**. As AI systems increasingly influence critical decisions in areas like hiring, healthcare, and criminal justice, ensuring **equitable outcomes** for all individuals and groups is paramount. This course addresses the urgent need for robust methodologies and metrics to detect, measure, and mitigate algorithmic bias, fostering the development and deployment of **responsible AI** solutions that uphold societal values and build public trust.

Training Course on Fairness and Accountability in AI Systems is designed to equip participants with the essential knowledge and practical skills to navigate the complex landscape of **AI ethics**. We will delve into various forms of AI bias, explore cutting-edge **fairness metrics**, and investigate practical **bias mitigation techniques**. Through a blend of theoretical foundations and real-world **case studies**, attendees will learn to implement **accountable AI governance frameworks**, ensuring that their AI systems are not only efficient and innovative but also inherently fair and transparent, ultimately driving a future where AI serves humanity justly.

Programme Curriculum

- Adversarial Debiasing: Using adversarial networks to reduce discriminatory attributes.
- Case Study 6.1: Generative AI Bias: Mitigating gender and racial stereotypes in image generation.
- Fairness-Aware Machine Learning Algorithms: Introduction to specialized algorithms.

Module 7: Explainable AI (XAI) and Transparency

- The Importance of Explainability: Why transparency is crucial for trust and accountability.
- Local vs. Global Explanations: Understanding individual prediction vs. overall model behavior.
- XAI Techniques: LIME, SHAP, Partial Dependence Plots, Feature Importance.

- Case Study 7.1: AI in Medical Diagnosis: Explaining complex AI decisions to clinicians and patients.
- Communicating AI Explanations Effectively: Tailoring explanations for diverse audiences.

Module 8: Implementing Accountability in AI Systems

- Defining Accountability in AI: Responsibility, auditability, recourse, and redress.
- Accountability Frameworks: Principles-based vs. prescriptive approaches.
- Role of Human Oversight: Ensuring human control and intervention in AI decision-making.
- Case Study 8.1: Autonomous Vehicles and Accountability: Determining liability in AI-driven accidents.
- Designing for Recourse: Mechanisms for challenging and correcting unfair AI outcomes.

Module 9: AI Governance and Ethical Guidelines

- Developing an Organizational AI Ethics Strategy: Policies, committees, and ethical review boards.
- Establishing AI Governance Frameworks: Integrating ethics into existing governance structures.
- Role of AI Ethics Charters and Principles: Guiding responsible AI development.
- Case Study 9.1: IBM's AI Ethics Board: Examining a large corporation's approach to AI governance.
- Cross-functional Collaboration: Engaging legal, compliance, technical, and business teams.

Module 10: Legal and Regulatory Landscape of AI Ethics

- Overview of Global AI Regulations: GDPR, CCPA, and emerging AI-specific laws (e.g., EU AI Act).
- Impact of Anti-Discrimination Laws on AI: How existing laws apply to algorithmic decision-making.
- Data Privacy and AI: Ensuring ethical data handling and compliance.
- Case Study 10.1: Algorithmic Discrimination Lawsuits: Learning from past legal challenges.
- Future Trends in AI Regulation: Anticipating upcoming compliance requirements.

Module 11: Auditing and Monitoring for Fairness and Accountability

- Fairness Audits: Methodologies for systematic evaluation of AI systems for bias.
- Continuous Monitoring: Techniques for detecting drift and emerging biases in deployed AI.
- Bias Reporting and Remediation: Establishing processes for identifying and addressing issues.
- Case Study 11.1: Algorithmic Impact Assessments (AIAs): Conducting ethical assessments before deployment.
- Tools for AI Model Monitoring: Utilizing platforms for ongoing performance and fairness checks.

Module 12: Building a Culture of Ethical AI

- Ethical AI Training and Education: Fostering awareness and competency across the organization.
- Incentivizing Responsible AI Practices: Aligning ethical goals with performance metrics.
- Promoting Diversity and Inclusion in AI Teams: Impact on bias mitigation.
- Case Study 12.1: Google's Ethical AI Challenges: Learning from internal dissent and initiatives.
- Leadership Role in Championing Ethical AI: Driving top-down commitment.

Module 13: Practical Implementation of Fairness and Accountability

- Fairness by Design: Integrating ethical considerations from the initial stages of AI development.
- Developing an AI Ethics Checklist: A practical guide for project teams.
- Risk Assessment for AI Systems: Identifying and prioritizing ethical risks.
- Case Study 13.1: Developing a Fair Lending AI System: A step-by-step approach to ethical design.
- Best Practices for Deploying Responsible AI: Checklist for secure and ethical deployment.

Module 14: Advanced Topics in AI Fairness Research

- Causal Fairness: Understanding causal relationships and their role in fairness.
- Fairness in Federated Learning and Privacy-Preserving AI: Challenges and solutions.
- Beyond Individual and Group Fairness: Exploring concepts like counterfactual fairness.
- Case Study 14.1: Debiasing Large Language Models (LLMs): Addressing societal biases in generative AI.
- Future Directions in AI Fairness Research: Open problems and emerging solutions.

Module 15: AI for Social Good and Equitable Outcomes

- Leveraging AI to Reduce Disparities: Applications of AI in healthcare, education, and social services.
- AI for Sustainable Development Goals (SDGs): Ethical deployment for global challenges.
- Ensuring Equitable Access to AI: Addressing the digital divide and technological inequalities.
- Case Study 15.1: AI in Disaster Response with Equity Considerations: Fair resource allocation during crises.
- The Future of Responsible AI: Collaborative efforts for a just and inclusive AI-powered world.

Training Methodology

This training course will employ a highly interactive and practical methodology, combining:

- **Expert-led Lectures and Discussions:** Engaging presentations by industry practitioners and academics.
- **Hands-on Workshops and Coding Labs:** Practical exercises using open-source tools (e.g., Python libraries for fairness metrics, bias mitigation).
- **Real-world Case Study Analysis:** In-depth examination of historical and contemporary examples of AI fairness challenges and solutions.
- **Group Activities and Collaborative Problem-Solving:** Fostering peer learning and diverse perspectives.
- **Interactive Q&A Sessions:** Opportunities for direct engagement with instructors and expert insights.
- **Guest Speaker Sessions:** Insights from leading AI ethicists, policymakers, and industry innovators.
- **Practical Frameworks and Templates:** Provision of actionable tools for immediate application in participants' organizations.

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	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

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

Register Online Now

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