

Fairness and Explainability in AI Systems

COURSE OVERVIEW

This course explores the critical concepts of fairness, transparency, and explainability in AI systems. The course includes practical tools and frameworks for measuring fairness and generating model explanations, equipping participants with the skills to build ethical and trustworthy AI applications. Participants will learn how biases are introduced into algorithms, how to detect and reduce them, and how to build models that are interpretable and accountable.

WHO SHOULD ATTEND?

This course is designed for data scientists, AI/ML engineers, technical leads, compliance officers, product managers, researchers, and policy advisors. All categories of professionals working on AI systems where transparency, accountability, and fairness are required or regulated, such as in finance, healthcare, human resources, or government, will find the course as a useful guide in their operations.

COURSE OUTCOMES

Delegates will gain the knowledge and skills to:

- Know about the sources and impact of bias in AI systems.
- Apply fairness metrics to evaluate model behavior.
- Use tools to explain model predictions and decisions.
- Design AI systems with transparency and ethical alignment.
- Communicate AI decisions clearly to non-technical audiences.
- Build trust in AI systems through interpretable design.
- Align model development with regulatory and ethical standards.

KEY COURSE HIGHLIGHTS

At the end of the course, you will understand;

- An introduction to fairness in AI: concepts and challenges.
- Types of bias: data, algorithmic, and societal.
- Fairness metrics: demographic parity, equal opportunity, etc.
- Explainable AI (XAI) tools: SHAP, LIME, and interpretable models.
- Trade-offs between performance, fairness, and interpretability.
- Regulatory compliance and ethical frameworks.
- Case studies in healthcare, finance, hiring, and law.
- Hands-on labs using Python-based fairness and XAI toolkits.
- Best practices for inclusive and responsible AI design.
- Communicating AI outcomes and decisions ethically.

All our courses are dual-certificate courses. At the end of the training, the delegates will receive two certificates.

1. A GTC end-of-course certificate
2. Continuing Professional Development (CPD) Certificate of completion with earned credits awarded