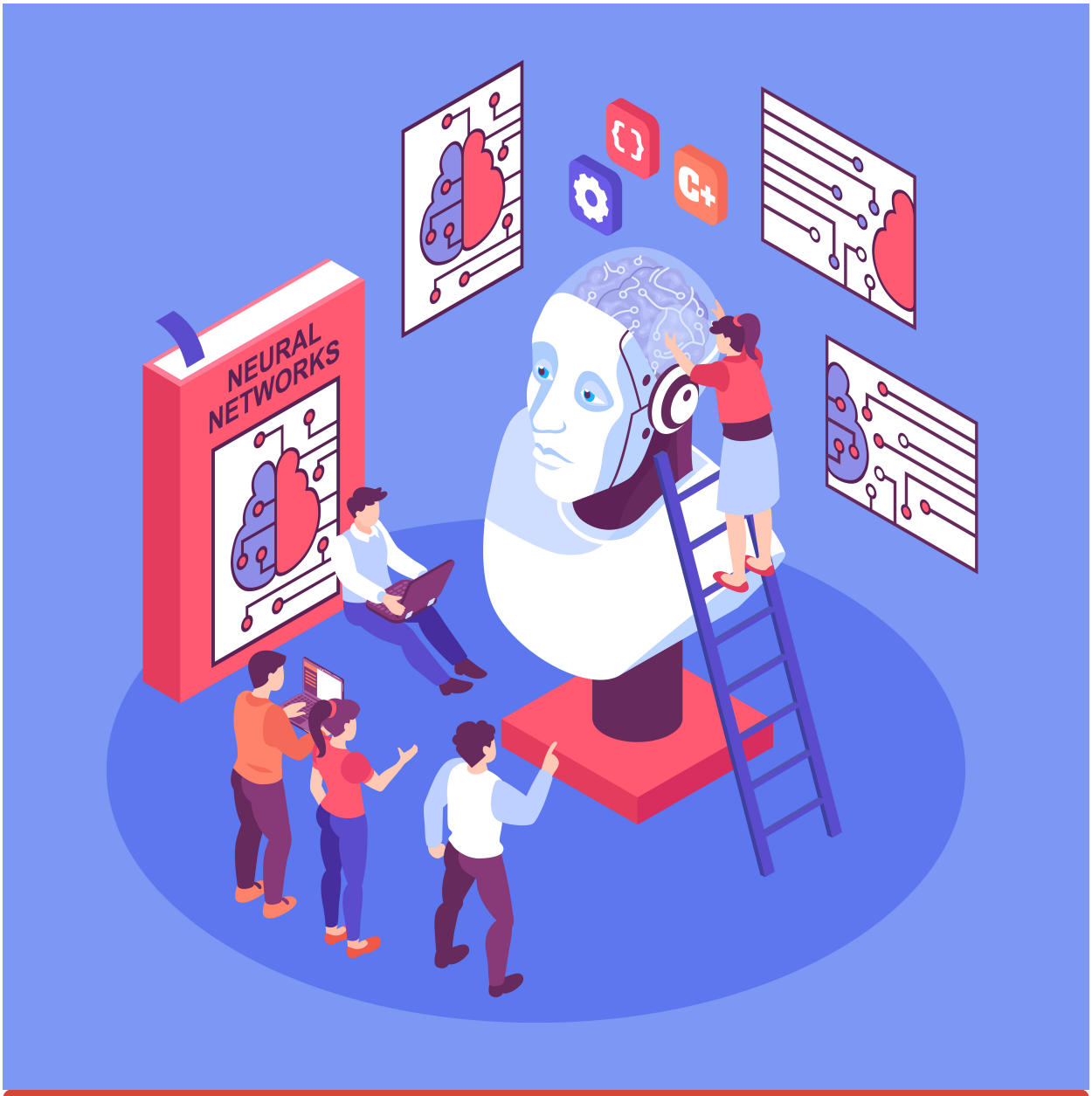

أخلاقيات الذكاء الاصطناعي والقدرة على التفسير: دليل الإبلاغ الآمن

المدة: 5 يوم

اللغة: ar

كود الكورس: PI2-111

هدف الكورس

بنهاية هذه الدورة، سيكون المشاركون قادرين على:

- فهم القضايا الأخلاقية الرئيسية في تطوير ونشر الذكاء الاصطناعي.
- تحديد المخاطر المتعلقة بالتحيز والتمييز والخصوصية.
- تطبيق الأطر والأدوات لتقييم ممارسات الذكاء الاصطناعي الأخلاقية.
- استكشاف الأساليب والأدوات لتحقيق الشفافية في نماذج الذكاء الاصطناعي.
- موازنة أداء النموذج مع الشفافية وقابلية التفسير.
- التنقل في حوكمة الذكاء الاصطناعي والامتثال والمعايير التنظيمية.
- بناء أو تقييم أنظمة ذكاء اصطناعي تكون مسؤولة اجتماعياً وقابلة للتدقيق.

الجمهور

هذه الدورة مثالية لـ:

- علماء البيانات، مهندسي الذكاء الاصطناعي، وممارسي تعلم الآلة.
- مديري المنتجات والقادة التقنيين الذين يطورون خدمات تعتمد على الذكاء الاصطناعي.
- مسؤولي الامتثال والمحترفين القانونيين في المؤسسات التقنية.
- صناع السياسات والمنظمين المشرفين على حوكمة الذكاء الاصطناعي.
- الأكاديميين والباحثين في مجالات الذكاء الاصطناعي، الأخلاقيات، أو الحقوق الرقمية.
- ممثلي المنظمات غير الحكومية والمدافعين المهتمين بعدالة الذكاء الاصطناعي.
- مستشاري تكنولوجيا المعلومات الذين ينفذون حلول الذكاء الاصطناعي في المجالات الحساسة.

منهجية التدريب

يستخدم هذا الدورة مزيجًا من الجلسات التي يقودها المدربون، ومناقشات حول المعضلات الأخلاقية، وتحليل الحالات الواقعية، وتمارين عملية لفهم الشرح. سيستكشف المشاركون أدوات تدقيق الذكاء الاصطناعي، ويقارنون بين النماذج الغامضة والنماذج القابلة للتفسير، ويقيّمون دراسات الحالة من خلال الأطر الأخلاقية. تعزز الدورة التفكير النقدي والحوار والتطبيق العملي.

الملخص

مع استمرار تأثير الذكاء الاصطناعي على اتخاذ القرارات في مجالات الرعاية الصحية والتمويل والعدالة وغيرها، أصبحت الأسئلة المتعلقة بالعدالة والمساءلة والشفافية في طليعة تطوير الذكاء الاصطناعي. يقدم هذا المقرر التحديات الأخلاقية التي تطرحها أنظمة الذكاء الاصطناعي ويركز على الذكاء الاصطناعي القابل للتفسير (XAI) كعنصر حيوي للثقة والامتثال والابتكار المسؤول.

سيستكشف المشاركون تقاطع الأخلاق والتنظيم والحلول التقنية المصممة لجعل أنظمة الذكاء الاصطناعي مفهومة وعادلة. يقدم المقرر توازناً بين الأسس المفاهيمية ودراسات الحالة الواقعية والأدوات العملية لتقييم وتحسين شفافية نماذج الذكاء الاصطناعي، خاصة في البيئات ذات المخاطر العالية أو الخاضعة للتنظيم.

محتوى الكورس والمخطط الزمني

Section 1: Foundations of AI Ethics

- .Why AI ethics matters: trust, safety, and social responsibility
- .Key ethical principles: fairness, accountability, transparency, autonomy
- .Common risks: bias, discrimination, opacity, surveillance
- .Real-world case studies: when AI fails or harms
- .Ethical frameworks: consequentialism, deontology, virtue ethics in AI
- .Stakeholder impact and ethical design considerations
- .Role of human oversight and moral reasoning

Section 2: Understanding Bias and Fairness in AI

- .(How bias enters AI systems (data, design, deployment
- .Types of bias: historical, algorithmic, selection, measurement
- .Evaluating fairness: individual vs. group fairness metrics
- .(.Tools for bias detection and mitigation (Fairlearn, AI Fairness 360, etc
- Fairness trade-offs: performance vs. equity
- .Inclusive dataset design and demographic parity
- .Guidelines for equitable model development

Section 3: Explainable AI (XAI) Principles and Techniques

- What is XAI and why it matters.
- Black-box vs. white-box models: strengths and limitations.
- Post-hoc explanation methods: SHAP, LIME, counterfactuals.
- Interpretable models: decision trees, rule-based systems, linear models.
- Domain-specific challenges in explainability (e.g., healthcare, finance).
- Communicating AI decisions to non-technical stakeholders.
- Hands-on walkthrough: explaining a black-box classifier.

Section 4: Governance, Compliance, and Regulation

- Overview of global AI regulations and guidelines (EU AI Act, OECD, UNESCO).
- Industry standards: ISO/IEC 24028, NIST AI Risk Management Framework.
- Ethics review boards, model documentation, and audit trails.
- AI in high-risk sectors: health, law, public services.
- Data protection, consent, and the right to explanation.
- Building internal governance structures for ethical AI.
- Preparing for regulatory audits and external evaluations.

Section 5: Designing Ethical and Transparent AI Systems

- Ethical design thinking in AI product development.
- Balancing accuracy, explainability, and user trust.
- Human-in-the-loop systems and oversight protocols.
- Transparency by design: UI, feedback, documentation.
- Building ethical AI cultures inside organizations.
- Scenario planning for ethical decision-making.
- Action planning: applying XAI and ethics to your projects.

تفاصيل الشهادة

عند إتمام هذه الدورة التدريبية بنجاح، سيحصل المشاركون على شهادة إتمام التدريب من Holistique Training. وبالنسبة للذين يحضرون ويكملون الدورة التدريبية عبر الإنترنت، سيتم تزويدهم بشهادة إلكترونية (e-Certificate) من Holistique Training.

شهادات Holistique Training معتمدة من المجلس البريطاني للتقييم المهني (BAC) وخدمة اعتماد التطوير المهني المستمر (CPD)، كما أنها معتمدة وفق معايير ISO 9001 وISO 21001 وISO 29993.

لهذه الدورة من خلال شهاداتنا، وستظهر هذه النقاط على شهادة إتمام التدريب من Holistique Training. يتم منح نقاط التطوير المهني المستمر (CPD) واحدة عن كل ساعة CPD يتم منح نقطة، CPD ووفقًا لمعايير خدمة اعتماد.

.لأي دورة واحدة نقدمها حاليًا CPD حضور في الدورة. ويمكن المطالبة بحد أقصى قدره 50 نقطة

التصنيفات

الذكاء الاصطناعي وإدارة البيانات, تطبيقات تكنولوجيا المعلومات والكمبيوتر, التكنولوجيا

مقالات ذات صلة

Top 15 Skills Every Data Scientist Needs in 2025

Discover the top 15 data scientist skills you need to succeed in 2025, including technical, .communication, cloud, and soft skills—plus expert tips on how to master them