الاصطناعي الدورة التدريبية: إدارة انحراف النموذج وقابلية التفسير في أنظمة الذكاء

يونيو - ٠٣ يوليو ٢٠٢٦ ٢٩

القاهرة - *

(للشخص الواحد) ٤١٠٠ €

Ref: #AI3042_559946

مقدمة الدورة التدريبية / لمحة عامة:

الذكاء الاصطناعي، وهي التدريبية المتخصصة حول إدارة انحراف النموذج يقدم BIG BEN Training Center هذه الدورة ومديري المنتجات، ومسؤولي الامتثال، والباحثين مصممة لعلماء البيانات، ومهندسي التعلم الآلي، وقابلية التفسير في أنظمة في البيئات الإنتاجية، والتعلم الآلي (.(Machine Learning مع تزايد نشر (AI الذين يعملون في مجال الذكاء الاصطناعي (وفهم كيفية اتخاذها للقرارات. تُعرف هذه يبرز تحدي أساسي يتمثل في الحفاظ على أدائها بمرور نماذج الذكاء الاصطناعي .(Box Problem والتحيز (Bias)، ومشكلة "الصندوق الأسود" (Black التحديات بانحراف النموذج (Model Drift)، الوقت، الاصطناعي**، واكتشاف انحراف النموذج، بالمعرفة والمهارات اللازمة لمراقبة أداء نماذج تهدف الدورة إلى تزويد المشاركين (Model Lifecycle لفهم سلوك النماذج، وإدارة دورة حياتها (XAI) وتطبيق تقنيات الذكاء الاصطناعي القابل للتفسير الذكاء (Concept Drift)، وتقنيات الكشف الدورة مفاهيم انحراف البيانات (Data Drift)، بشكل فعال. ستغطي (Management النماذج**. كما ستركز على أهمية قابلية التفسير في عن الانحراف، والاستراتيجيات المتبعة لإعادة تدريب وانحراف المفهوم لأداء النماذج في الوقت الرائدة مثل LIME و .SHAP سيتعلم المشاركون كيفية بناء الثقة والامتثال، واستعراض أدوات XAI وتطوير حلول لتخفيف تأثير الانحراف**، وضمان الفعلي**، وتفسير التنبؤات والقرارات الصادرة عنها، تصميم أنظمة مراقبة تهدف الدورة إلى تمكين

وموثوقة، وتعزيز الثقة في تقنيات الذكاء المختصين من بناء وصيانة أنظمة ذكاء اصطناعي قوية الشفافية والمساءلة.

رودين(Cynthia Rudin) ، والأخلاقية**. نستلهم في هذه الدورة من أعمال الاصطناعي، والالتزام بالمتطلبات التنظيمية

للتفسير بطبيعتها لتعزيز الثقة والمساءلة. ، التي تركز على تطوير نماذج قابلةCynthia Rudin البروفيسور سينثيا

الفئات المستهدفة / هذه الدورة التدريبية مناسبة لـ:

- علماء البيانات.
- مهندسي التعلم الآلي.
- مهندسي العمليات والتشغيل (MLOps Engineers).
- مديري المنتجات.
- مسؤولي الامتثال والحوكمة.
- محللي جودة البيانات.
- الباحثين في الذكاء الاصطناعي.
- المدققين التقنيين.
- مديري المخاطر التقنية.
- المهتمين بالحفاظ على أداء نماذج الذكاء الاصطناعي.

القطاعات والصناعات المستهدفة:

- الخدمات المالية (كشف الاحتيال، تقييم الائتمان).
- الرعاية الصحية (التشخيص، التنبؤ بالأمراض).
- البيع بالتجزئة والتجارة الإلكترونية.
- الاتصالات.
- التصنيع (الصيانة التنبؤية).
- التكنولوجيا والبرمجيات.
- القطاع الحكومي (أنظمة دعم القرار).
- السيارات ذاتية القيادة.
- الأمن السيبراني.
- الاستشارات التقنية.

الأقسام المؤسسية المستهدفة:

- قسم علوم البيانات.
- قسم التعلم الآلي والذكاء الاصطناعي.
- قسم عمليات التعلم الآلي (MLOps).
- قسم إدارة المنتجات (المنتجات القائمة على الذكاء الاصطناعي).
- قسم الحوكمة والمخاطر والامتثال (GRC).
- قسم الجودة والتدقيق.
- قسم البحث والتطوير.
- قسم هندسة البيانات.
- قسم إدارة المخاطر.
- القسم القانوني.

أهداف الدورة التدريبية:

أتقن المهارات التالية: بنهاية هذه الدورة التدريبية، سيكون المتدرب قد

- فهم مفاهيم انحراف البيانات وانحراف المفهوم.
- المتقدمة. اكتشاف انحراف النموذج باستخدام تقنيات المراقبة
- في الإنتاج. تصميم أنظمة لمراقبة أداء نماذج الذكاء الاصطناعي
- (XAI). تطبيق تقنيات الذكاء الاصطناعي القابل للتفسير
- تفسير سلوك نماذج التعلم الآلي المعقدة.
- (Data Drift, Concept Drift, Model Decay) تحديد أسباب انحراف النموذج
- وضع استراتيجيات لإعادة تدريب النماذج وتحديثها.
- الاصطناعي. ضمان الشفافية والمساءلة في قرارات الذكاء
- التقنيين. توصيل رؤى قابلية التفسير لأصحاب المصلحة غير
- بناء أنظمة ذكاء اصطناعي قوية وموثوقة ومستدامة.

منهجية الدورة التدريبية:

والتطبيق العملي منهجية تدريبية تجمع بين الفهم النظري المتعمق يعتمد BIG BEN Training Center في هذه الدورة على الاصطناعي بكفاءة ومسؤولية في بيئات المكثف، بهدف تمكين المشاركين من إدارة أنظمة لـانحراف النموذج وقابلية التفسير والامتثال. تتبع ذلك أنواع الانحرافات التي تؤثر على أداء النماذج، الإنتاج. تشمل المنهجية محاضرات تفاعلية تستعرض الذكاء مراقبة لـأداء النماذج**، واكتشاف ورش عمل تطبيقية مكثفة حيث سيقوم المشاركون بـتصميم وأهمية XAI في بناء الثقة انحراف وSHAP) على سيناريوهات واقعية. سيتم التركيز على LIME الرائدة (مثل XAI الانحراف، وتطبيق أدوات وتنفيذ أنظمة الأداء وضمان العدالة. تتضمن النموذج في تطبيقات حقيقية، وكيفية تفسير سلوك دراسات حالة عملية تبرز كيفية التعامل مع تغذية متكاملة لـإدارة دورة حياة النموذج**، بما في ذلك الدورة جلسات عمل جماعي لـتطوير استراتيجيات النماذج لـتحسين هذا المجال الحيوي والمعقد. راجعة مفصلة ومنتظمة من المدربين الخبراء لضمان إعادة التدريب والتحقق. يتلقى المشاركون تطوير مهاراتهم في

خريطة المحتوى التدريبي (محاور الدورة التدريبية):

المراقبة. الوحدة الأولى: فهم انحراف النموذج وأهمية

- مقدمة إلى انحراف النموذج (Model Drift).
- انحراف المفهوم (Concept Drift) أنواع الانحراف: انحراف البيانات (Data Drift)،
- تأثير الانحراف على أداء نماذج الذكاء الاصطناعي.
- أهمية المراقبة المستمرة لأداء النماذج في الإنتاج.
- مراحل دورة حياة نموذج الذكاء الاصطناعي (MLOps).
- مؤشرات الأداء الرئيسية (KPIs) لمراقبة النموذج.
- تحديات إدارة النماذج في البيئات الديناميكية.

الوحدة الثانية: تقنيات الكشف عن انحراف البيانات والمفهوم.

- طرق الكشف عن انحراف البيانات:
- اختبارات التوزيع الإحصائي (Statistical Distribution Tests).
- قياس المسافة بين التوزيعات (Divergence Measures).
- مراقبة الميزات والافتراضات.
- طرق الكشف عن انحراف المفهوم:
- الكشف المبكر عن التغييرات في العلاقات بين البيانات والنتائج.
- المقاييس المستندة إلى الأداء (Performance-based metrics).
- أمثلة على أدوات الكشف عن الانحراف (Alibi Detect, evidently AI).
- تصميم لوحات تحكم لمراقبة الانحراف.

الوحدة الثالثة: قابلية التفسير (XAI) لفهم سلوك النموذج.

- مقدمة إلى الذكاء الاصطناعي القابل للتفسير (XAI).
- أهمية XAI في فهم الانحراف والتحيز.
- تقنيات التفسير المحلية (Local Interpretability): LIME.
- تقنيات التفسير المحلية (Local Interpretability): SHAP.
- تطبيق LIME و SHAP لتفسير التنبؤات الفردية.
- مقاييس أهمية الميزات (Feature Importance).
- التمييز بين الشفافية والقابلية للتفسير.

الوحدة الرابعة: استراتيجيات إدارة الانحراف وإعادة التدريب.

- لماذا ومتى يجب إعادة تدريب النماذج؟
- أنواع إعادة التدريب: إعادة التدريب اليدوي، إعادة التدريب الآلي.
- استراتيجيات اختيار البيانات لإعادة التدريب الفعال.
- التكيفي (Adaptive Learning)، التعلم المستمر (Continual Learning) والتعلم
- تخفيف تأثير الانحراف على الأداء.
- إدارة إصدارات النماذج (Model Versioning).
- بناء خطوط أنابيب (Pipelines) قوية لإعادة التدريب والنشر.

الوحدة الخامسة: الممارسات المتقدمة في XAI وحوكمة النماذج.

- تفسير الشبكات العصبية العميقة.
- العدالة (Fairness) والتحيز (Bias) في نماذج الذكاء الاصطناعي.
- أخلاقيات الذكاء الاصطناعي وانحراف النموذج.
- توصيل رؤى XAI لأصحاب المصلحة.
- حوكمة النموذج (Model Governance) والامتثال التنظيمي.
- المساءلة والتدقيق لأنظمة الذكاء الاصطناعي.
- الآفاق المستقبلية في إدارة انحراف النموذج وXAI.

الأسئلة المتكررة:

التسجيل في الدورة؟ ما هي المؤهلات أو المتطلبات اللازمة للمشاركين قبل

لا توجد شروط مسبقة.

الإجمالي لساعات الدورة التدريبية؟ كم تستغرق مدة الجلسة اليومية، وما هو العدد

المدة إلى ٢٥٢٠- بمعدل يومي يتراوح بين ٤ إلى ٥ ساعات، تشمل فترات تمتد هذه الدورة التدريبية على مدار خمسة أيام، ساعة تدريبية. راحة وأنشطة تفاعلية، ليصل إجمالي

سؤال للتأمل:

من خلال تقنيات في التطبيقات الحيوية، هل يمكننا فعلاً تحقيق فهم مع تزايد تعقيد نماذج الذكاء الاصطناعي وانتشارها التي يمكننا بلوغها مع الأنظمة قابلية التفسير الحالية، أم أن هناك حدوداً جوهرية كامل وموثوق لسلوك هذه النماذج المعقدة بشكل متزايد؟ لمدى الشفافية

ما الذي يميز هذه الدورة عن غيرها من الدورات؟

هو دمج تقنيات لإدارة تحديين أساسيين في الذكاء الاصطناعي: انحراف تتميز هذه الدورة بتقديمها نهجاً متكاملاً وعملياً الاصطناعي القابل للتفسير(XAI)، الكشف عن الانحراف واستراتيجيات إعادة التدريب مع النموذج وقابلية التفسير. ما يميزنا تصميم أنظمة النماذج وكيفية عملها داخلياً. نغطي الجوانب مما يتيح للمشاركين فهم أعمق لسبب تغير أداء أدوات الذكاء تمكين المختصين من بناء وصيانة مراقبة فعالة وتطبيق XAI على سيناريوهات واقعية. النظرية والتطبيقية، مع التركيز على والمساءلة، والتخفيف من المخاطر المرتبطة بانحراف أنظمة ذكاء اصطناعي قوية وموثوقة، وضمان الشفافية الدورة تركز على بيئة الإنتاج. لإتقان إدارة دورة حياة نماذج الذكاء الاصطناعي في النموذج**، مما يجعلها ضرورية لأي محترف يسعى