



عمليات تعلم الآلة ML Ops و LLM Ops: نشر النماذج ومراقبتها وتحسينها في بيئة الإنتاج



عمليات تعلم الآلة MLOps و LLMOps: نشر النماذج ومراقبتها وتحسينها في بيئة الإنتاج

مقدمة:

تسد عمليات تعلم الآلة MLOps وعمليات النماذج اللغوية الكبيرة LLMOps الفجوة بين نموذج يحقق نتائج جيدة في دفتر التجارب وخدمة تبقى دقيقة وسريعة ومعقولة الكلفة بعد الإطلاق. فكثير من الفرق ما زالت تعيد التدريب يدويا، ولا تعرف أي مجموعة بيانات أنتجت النموذج العامل، ولا تكتشف الانجراف إلا بعد شكاوى المستخدمين، وتفاجأ بارتفاع فواتير الرموز في تطبيقات النماذج اللغوية. تقدم هذه الدورة من كور كونسبت للمهندسين منهجية محايدة تجاه الأدوات لإدارة إصدارات النماذج التقليدية وتطبيقات النماذج اللغوية ونشرها وتقديمها ورصدها والتراجع عنها. ويصمم المشاركون مخطط مسار MLOps و LLMOps مع لوحة مراقبة لخدمة ذكاء اصطناعي محل دراسة حالة.

أهداف الدورة:

- تقييم عملية الإصدار الحالية لنماذج تعلم الآلة والنماذج اللغوية لدى الفريق مقابل مستويات نضج MLOps ومقياس Score ML Test لقياس جاهزية الإنتاج
- إدارة إصدارات مجموعات البيانات والسمات والموجهات وملفات النماذج بحيث يمكن تتبع أي تنبؤ في الإنتاج إلى جولة التدريب التي أنتجته
- بناء مسارات تكامل ونشر مستمرين CI/CD وتدريب مستمر تختبر البيانات وتعيد التدريب وتسجل النماذج وترقيها عبر بيئات مرحلية
- اختيار نمط التقديم المناسب بين الدفعي والفوري والتدفقي والطرفي وضبط التوسع التلقائي والإطلاق التدريجي والتراجع لكل منها
- كشف انجراف البيانات وانجراف المفهوم وتراجع جودة النماذج اللغوية باختبارات إحصائية وأطر تقييم وحدود تنبيه
- ضبط إنفاق الرموز ومدة الاستجابة والمخرجات غير الآمنة في خدمات النماذج اللغوية عبر التخزين المؤقت وتوجيه النماذج وحواجز الحماية والتتبع

الفئة المستهدفة:

- مهندسو تعلم الآلة الذين يحزمون النماذج التنبؤية وينشرونها ويصنعونها خلف واجهات برمجية في الإنتاج
- علماء البيانات الذين يسلمون النماذج المدربة لفرق الهندسة ويحتاجون إلى إبقائها دقيقة بعد الإطلاق
- مهندسو المنصات والبنية التحتية الذين يشغلون موارد الحوسبة والحاويات والمسارات التي تستضيف أحمال الذكاء الاصطناعي
- مطورو تطبيقات النماذج اللغوية الكبيرة المسؤولون عن كلفة الميزات التوليدية وسرعتها وجودتها إجاباتها
- مهندسو موثوقية المواقع والمراقبة الذين يضمنون خدمات الذكاء الاصطناعي إلى المناوبات والتنبهات وإجراءات الحوادث

محاور الدورة:

اليوم 1: دورة حياة تعلم الآلة والنماذج اللغوية في الإنتاج وخط أساس النضج

- الفجوة بين دفتر التجارب والخدمة: الدين التقني الخفي في أنظمة تعلم الآلة الإنتاجية
- دورة حياة تعلم الآلة التقليدي مقابل دورة حياة تطبيقات النماذج اللغوية: مراحل التدريب وصياغة الموجهات والاسترجاع
- مستويات نضج MLOps: العملية اليدوية ومسار التدريب المؤتمت والأتمتة الكاملة للتكامل والنشر المستمرين
- معيار ML Test Score لتقييم جاهزية البيانات والنموذج والبنية التحتية والمراقبة
- تدقيق خط الأساس لعملية الإصدار لدى الفريق وخريطة النضج المستهدف

اليوم 2: قابلية إعادة الإنتاج وإدارة الإصدارات وبنية المسارات

- إدارة إصدارات البيانات والسمات والشفرة بالبصمات الرقمية للمحتوى ولقطات مجموعات البيانات
- تتبع التجارب: بيانات الجولات الوصفية والمعاملات الفائقة والمقاييس وسلسلة نسب الملفات
- تصميم مخزن السمات: المخزن غير المتصل والمخزن الفوري والتباين بين التدريب والتقديم
- مراحل سجل النماذج وتوابع النماذج وبوابات اعتماد الترقية
- تنسيق سير العمل بمسارات الرسوم البيانية الموجهة غير الدورية والخطوات المحواة

اليوم 3: التسليم المستمر وتقديم النماذج وتوسيعها

- التكامل والنشر المستمران لتعلم الآلة: اختبارات التحقق من البيانات وبوابات جودة النموذج وملفات البناء
- محفزات التدريب المستمر: الجدولة وحجم البيانات وتراجع الأداء
- أنماط النشر: التقييم الدفعي وواجهات الاستدلال الفوري والتدفق والنشر على الأجهزة الطرفية
- بيئات تشغيل تقديم النماذج وتجميع الطلبات على GPU والتكميم والتوسع الأفقي التلقائي
- النشر الظلي والإطلاق التدريجي واختبار A/B والتراجع المؤتمت

اليوم 4: مراقبة الانجراف وتقييم النماذج اللغوية وضبط الكلفة

- كشف انجراف البيانات وانجراف المفهوم باختبارات النوافذ DDM وADWIN وKSWIN
- إطار تقييم النماذج اللغوية: مجموعات الأسئلة المرجعية وقياس الاستناد إلى المصادر وملاءمة الإجابة
- إدارة إصدارات الموجهات وضبط الاسترجاع وتحسين تقسيم المقاطع وإعادة الترتيب في RAG التوليد المعزز بالاسترجاع
- حواجز الحماية وتصفية المخرجات: السمية وحجب البيانات الشخصية وفحص المخطط وقواعد الرفض
- ميزانيات الرموز والتخزين المؤقت الدلالي وتوجيه النماذج وأهداف مدة الاستجابة بالمئينات

اليوم 5: قابلية الرصد والاستجابة للحوادث ومختبر مخطط المسار

- تتبع الموزع لاستدعاءات الاسترجاع والموجهات والنماذج بقياسات تفصيلية لكل مرحلة من مراحل الطلب
- دليل تشغيل حوادث خدمات الذكاء الاصطناعي: مستويات الخطورة والتراجع إلى آخر نموذج سليم ومراجعة ما بعد الحادث
- نقاط ربط الحوكمة: بطاقات النماذج وسجلات النسب وسجلات اعتماد التغيير داخل المسار
- بناء مختبري: لوحة مراقبة للانجراف والجودة والكلفة ومدة الاستجابة لخدمة ذكاء اصطناعي محل دراسة حالة
- عرض مخطط مسار MLOps وLLMOps ومراجعة التصميم مع المشاركين

المهارات التي ستكتسبها:

- تتبع نسب النماذج
- تنسيق مسارات تعلم الآلة
- تصميم التدريب المستمر
- تحسين تقديم الاستدلال
- هندسة كشف الانجراف
- تقييم جودة النماذج اللغوية
- حوكمة كلفة الرموز
- رصد خدمات الذكاء الاصطناعي

لماذا تحضر هذه الدورة:

- العودة بمخطط مسار MLOps وLLMOps وتصميم لوحة مراقبة جاهزة لخدمة ذكاء اصطناعي محل دراسة حالة
- التوقف عن اكتشاف تراجع النماذج من شكاوى المستخدمين بتفعيل تنبيهات الانجراف والجودة قبل الإطلاق
- خفض الفواتير المفاجئة لخدمات النماذج اللغوية وبطء الاستجابة عبر التخزين المؤقت والتوجيه وحدود الميزانية لكل نقطة خدمة
- مقارنة ممارسات الإصدار والتقديم والمناوبة مع مهندسين يشغلون الذكاء الاصطناعي في المال والتجزئة والطاقة والاتصالات والخدمات العامة

الخاتمة:

لا تحافظ النماذج وتطبيقات النماذج اللغوية على قيمتها إلا حين تمر بياناتها وشيفرتها وموجهاتها وملفاتها عبر مسار واحد قابل للتتبع والمراقبة. تنتقل الدورة من دورة الحياة وخط أساس النضج، إلى إدارة الإصدارات وتتبع التجارب وسجلات النماذج وتنسيق المسارات، ثم التسليم المستمر والتقديم والتوسع، فكشف الانجراف وتقييم النماذج اللغوية وحواجز الحماية وضبط الكلفة. ويغطي اليوم الأخير التتبع وأدلة الحوادث ونقاط ربط الحوكمة، ويبني مخطط مسار MLOps و LLMOps مع لوحة مراقبة لخدمة ذكاء اصطناعي.