



دورة الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML للأجهزة الصناعية وإنترنت الأشياء



دورة الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML للأجهزة الصناعية وإنترنت الأشياء

مقدمة:

دورة الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML للأجهزة الصناعية وإنترنت الأشياء دورة تمتد أسبوع، موجهة لفرق الأنظمة المدمجة والأتمتة والأجهزة الدقيقة وحلول إنترنت الأشياء، وتنتهي بإعداد خطة نشر الذكاء الاصطناعي على الحافة لمصنع أو أسطول أجهزة افتراضي. فكثير من المنشآت الصناعية تريد تشغيل نماذج تعلم الآلة داخل الكاميرات ومستشعرات الاهتزاز والبوابات، غير أن النماذج المصممة للحواد تم تجاوز ذاكرة المتحكمات الدقيقة أو تستنزف البطاريات أو تتوقف حين ينقطع اتصال الشبكة. ويعمل المشاركون أصلا على بناء الأجهزة الميدانية أو دمجها، ويتدربون في مختبر عملي على لوحات المتحكمات الدقيقة ووحدات الحافة. وتقدم كور كونسبت هذه الدورة.

أهداف الدورة:

- تقييم حالات الاستخدام الصناعية لتحديد موضع الاستدلال على الجهاز أو البوابة أو السحابة وفق معايير زمن الاستجابة وعرض النطاق والاتصال
- اختيار المتحكمات الدقيقة ووحدات المعالجة العصبية NPU ووحدات GPU وفق ميزانيات الذاكرة والإنتاجية والطاقة لكل حمل عمل
- ضغط النماذج المدربة بالتكميم إلى INT8 والتقليم ونقل المعرفة من نموذج معلم إلى نموذج طالب مع تتبع فاقد الدقة
- نشر نماذج الرؤية وتمييز الكلمات المفتاحية وتصنيف الاهتزاز باستخدام TensorFlow Lite Micro و ONNX Runtime على اللوحات المستهدفة
- بناء كشف الشذوذ على الجهاز لإشارات مقياس التسارع والإشارات الصوتية مع ملفات طاقة تعتمد التشغيل الدوري
- تخطيط توزيع النماذج عبر الهواء ومراقبة الأسطول لرصد الأجهزة المتراجعة قبل انخفاض جودة مخرجاتها

الفئة المستهدفة:

- فرق هندسة الأنظمة المدمجة التي تكتب البرامج الثابتة للمستشعرات ووحدات التحكم والكاميرات الذكية
- فرق الأتمتة والأجهزة الدقيقة المسؤولة عن مراقبة حالة المعدات في المصانع
- فرق حلول إنترنت الأشياء التي تربط البوابات والأجهزة الميدانية بمنصات بيانات المصنع
- فرق الرؤية الآلية المسؤولة عن كاميرات الفحص على خطوط الإنتاج والتعبئة
- فرق هندسة الاعتمادية التي تضيف الاستشعار والتحليلات إلى المعدات الدوارة

محاوِر الدورة:

اليوم 1: مشهد الذكاء الاصطناعي على الحافة وتأهيل حالات الاستخدام

- مقارنة مستويات الاستدلال بين السحابة والبوابة والجهاز
- معايير زمن الاستجابة وعرض النطاق والاتصال لتحديد التشغيل على الجهاز
- فئات أجهزة TinyML من المتحكمات الدقيقة إلى وحدات الحافة
- مصفوفة فرز حالات الاستخدام الصناعية للكاميرات والمستشعرات
- حصر الوضع الحالي لأصول المصنع وتدفقات بياناته

اليوم 2: معمارية عتاد الحافة وبيئات تشغيل الاستدلال

- خريطة ذاكرة المتحكم الدقيق بين ذاكرة الفلاش و SRAM ومساحة الموترات
- اختيار وحدات NPU و VPU و GPU باستخدام تصنيفات TOPS
- إعداد مفسر TensorFlow Lite Micro ومحلل العمليات
- مقارنة ONNX Runtime ومزودي التنفيذ في حزم تطوير المصنعين
- نتائج اختبارات MLPerf Tiny المرجعية لتصفية خيارات العتاد

اليوم 3: ضغط النماذج وبناء الاستدلال على الجهاز

- التكميم إلى INT8 بعد التدريب باستخدام عينات معايرة ممثلة
- إعادة التدريب الواعي بالتكميم لاستعادة دقة المصنف
- التقليم الهيكلي لتقليص حجم النموذج في الفلاش والذاكرة العشوائية
- نقل المعرفة Knowledge Distillation من شبكة معلمة إلى نموذج طالب مدمج
- تحميل مصنف الكلمات المفتاحية والاهتزاز على لوحات المتحكمات الدقيقة

اليوم 4: الرؤية على الجهاز وكشف الشذوذ وتشغيل الأساطيل

- كشف العيوب بكاميرات منخفضة الدقة ضمن ميزانية معدل الإطارات
- تقييم الشذوذ بالمشفر التلقائي على إشارات التسارع والصوت
- قياس ميزانية الطاقة بالتشغيل الدوري وأنماط السكون
- توزيع النماذج عبر الهواء بطلاقات إطلاق مرحلية للأسطول
- تسجيل الثقة على الجهاز ومؤشرات الانحراف للوحات متابعة الأسطول

اليوم 5: مختبر على مصنع افتراضي وخطة نشر الذكاء الاصطناعي على الحافة

- عرض حالة المصنع لمواقع المضخات والسيور الناقلة والكاميرات
- قائمة العتاد المختصرة وميزانية الذاكرة لأجهزة الحالة
- جدول الموازنة بين دقة النموذج المضغوط وزمن استجابته واستهلاكه للطاقة
- قواعد تحديث الأسطول ومراقبته لإطلاق الحالة
- استكمال خطة نشر الذكاء الاصطناعي على الحافة ومراجعة التصميم بين المشاركين

المهارات التي ستكتسبها:

- تحليل موضع الاستدلال
- اختيار عتاد الحافة
- تكميم النماذج
- تقليم الشبكات العصبية
- نشر النماذج على المتحكمات الدقيقة
- كشف شذوذ الاهتزاز
- تصميم الاستدلال منخفض الطاقة
- إطلاق النماذج على الأساطيل

لماذا تحضر هذه الدورة:

- إعداد خطة نشر الذكاء الاصطناعي على الحافة لمصنع أو أسطول أجهزة افتراضي ورفعها إلى مدير الهندسة وقائد تقنيات التشغيل
- الحسم في موضع تشغيل النموذج بين المتحكم الدقيق ووحدة الحافة والبوابة والسحابة استناداً إلى أرقام مقيسة للذاكرة وزمن الاستجابة والطاقة
- تجنب تعثر المشاريع التجريبية ونفاد البطاريات وتوقف الأجهزة عند انقطاع شبكة المصنع عبر موازنة حجم النماذج مع العتاد قبل الشراء
- نقل مصفوفة فرز حالات الاستخدام وقالب قائمة العتاد وقائمة فحص الضغط إلى فرق الأنظمة المدمجة والأتمتة

الخاتمة:

بعد العودة إلى العمل يقدم المشارك خطة نشر الذكاء الاصطناعي على الحافة إلى مدير الهندسة وقائد تقنيات التشغيل، فيستندان إلى قائمة العتاد وميزانيات الذاكرة والطاقة ونتائج الضغط في تحديد الأجهزة التي تشتري والنماذج التي تشغل عليها أولاً، وتوجه قواعد تحديث الأسطول ومراقبته طريقة وصول الإصدارات الجديدة من النماذج إلى المستشعرات والكاميرات، وبعد أول إطلاق وفق الخطة، تراجع الوحدة دقة النماذج على الأجهزة مقارنة بنتائج المختبر، وعمر البطاريات، والتحديثات المتعثرة، والأجهزة التي تظهر انحرافاً.

الأسئلة الشائعة:

ما الذي يحتاجه المشارك قبل دورة الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML؟

يحتاج المشارك إلى قراءة شيفرة Python الأساسية، وفهم طريقة تقييم مصنف مدرب، وخبرة في العمل على لوحات المتحكمات الدقيقة أو Linux المدمج. ولا يشترط التعمق في نظرية الشبكات العصبية، ويفيده إحضار وصف لحالة استخدام واحدة لمستشعر أو كاميرا من موقعه.

بم يختلف الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML عن دورات التعلم العميق أو عمليات تعلم الآلة؟

تبدأ هذه الدورة من نموذج مدرب مسبقاً وتعمل على مواءمته مع عتاد محدود الموارد من حيث الذاكرة والطاقة والضغط وتحديث الأساطيل. أما دورات التعلم العميق فتركز على بناء الشبكات وتدريبها، وتعنى دورات عمليات تعلم الآلة بتشغيل النماذج خدمات سحابية أو في مراكز البيانات.

لماذا يهم التحكم في تعلم الآلة المدمج على أجهزة إنترنت الأشياء؟

يخزن التحكم الأوزان والتنشيطات أعداداً صحيحة بثمانية خانات ثنائية بدلاً من أعداد عشرية أكبر، فيصغر النموذج وتتمكن وحدات NPU والمتحكمات الدقيقة التي تعمل بالأعداد الصحيحة من تشغيله بسرعة وطاقته أقل، مع فاقد محدود في الدقة عادة.

ماذا يعود به المشارك من دورة الذكاء الاصطناعي على الحافة وتعلم الآلة المدمج TinyML؟

يعود المشارك بخطة نشر الذكاء الاصطناعي على الحافة لمصنع أو أسطول افتراضي، تضم مصفوفة فرز حالات الاستخدام وقائمة العتاد المختصرة وميزانيات الذاكرة والطاقة ونتائج الضغط وقواعد تحديث الأسطول ومراقبته، جاهزة للتكليف مع أجهزة جهته.