

الذكاء الاصطناعي بطريقة صحيحة: مع أساس قوي، يمكن للجميع تخطيط رحلة ناجحة - بما في ذلك أنت

كيف يضع نموذج الذكاء الاصطناعي "البيانات - التدريب - الاستدلال" من
شركة IBM مؤسستك على طريق النجاح على المدى الطويل



نظرة عامة

ماذا لو اكتشفت أن خلفية علوم البيانات ليست ضرورية لفهم أكثر الجوانب تعقيدًا في مؤسسة تدار بواسطة الذكاء الاصطناعي وتفسيرها والتصرف فيها؟

قبل الانغماس في التفاصيل، من الهام معرفة أن نموذج "البيانات-التدريب-الاستدلال (DTI)" ليس تدفق عمل خطي. إنما هو حلقة مستمرة تتكون من ثلاث مراحل تتفاعل في جميع الأوقات. ولأن العملية مستمرة، فإن الأفكار المستخرجة أكثر ثراءً وقيمة.

النتيجة؟ يمكن لأصحاب المصلحة الرئيسيين اتخاذ قرارات أكثر ذكاءً بشكل أسرع - وبنقطة أكبر.

كيف ستؤثر هذه المعرفة على العمليات والتطبيقات التجارية التي تعتمد عليها مؤسستك لتبقى قادرة على المنافسة؟

كانت المؤسسات التي تدار بواسطة الذكاء الاصطناعي بصورة تقليدية محور خبراء التحليل مع فهم عميق لنماذج البناء والتدريب. ولكن كل هذا يتغير مع استمرار المبادرات المستندة إلى الذكاء الاصطناعي في جميع أنحاء العمل. وفي طليعة هذا التحول إطار عمل قائم على القيمة لجهود الذكاء الاصطناعي. ويطلق عليه نموذج الذكاء الاصطناعي "البيانات-التدريب-الاستدلال (DTI)"، وشرح ذلك هو الهدف من هذه الورقة البحثية.

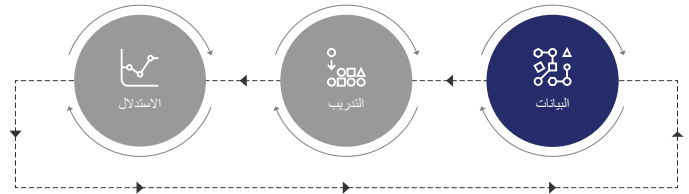
نموذج الذكاء الاصطناعي "البيانات-التدريب-الاستدلال (DTI)"



المرحلة 1: البيانات



”البيانات المشكوك فيها من حيث الجودة أو الصحة أو حتى الكمية سوف تنشئ نماذج ذكاء اصطناعي مشكوك فيها، مما يولد نتائج مشكوك فيها.“



يشير خبراء الذكاء الاصطناعي إلى مرحلة البيانات باعتبارها الأكثر استهلاكاً للوقت من المراحل الثلاث. كمية العمل المطلوبة لإعداد البيانات الحالية للتغذية - المصطلح الشائع لتحميل البيانات في نموذج الذكاء الاصطناعي للتدريب - هو مكثف.

كما يقول المثل القديم، أن أي تساؤل أو فرضية فارغة أو لا معنى لها، ستؤدي بالتالي إلى نتائج لا معنى لها. وبعد هذا الأمر أصدق من أي وقت مضى في عالم الذكاء الاصطناعي. بدايةً من البيانات المشكوك فيها من حيث الجودة أو الصحة أو حتى الكمية، ستنشئ نماذج الذكاء الاصطناعي المشكوك فيها التي تؤدي إلى نتائج مغلوطة بالنسبة إلى العمل. إن البدء دون إنشاء قاعدة بيانات قوية سوف يؤدي لإرسال مشاريع الذكاء الاصطناعي إلى المسار الخطأ قبل أن تبدأ.

هل كنت تعلم؟

أدت المساحات الشاسعة من البيانات التي تتخلل عالمنا، مقترنة بالقدرات المتزايدة بسرعة للبنية التحتية لخوادمنا، إلى إشعال ثورة الذكاء الاصطناعي في العقد الماضي.

تأتي البيانات المستخدمة في التدريب على الذكاء الاصطناعي من جميع أنواع المصادر المختلفة. قد تكون هذه مصادر مألوفة ومفهومة جيداً، مثل المبيعات السابقة أو أرقام العملاء من مستودع بيانات المؤسسة الحالي. قد تأتي أيضاً في الوقت الفعلي من مصادر مثل أجهزة إنترنت الأشياء على الحافة أو تدفقات بيانات الإنترنت الأخرى، مثل Twitter.

الحقائق الأربعة حول البيانات

أولاً، تأتي البيانات من العديد من المصادر المختلفة. قد يكون أي تنسيق غير معالج مثل النص أو الصورة أو الصوت أو القيم الرقمية الأولية. تقضي فرق علوم البيانات كميات هائلة من الوقت في جمع البيانات، ثم تنظيف هذه البيانات في تنسيقات صحيحة لتستهلكها اطرار الأعمال الموجودة في البيئة. وهذه خطوات هامة. يعد تحليل وتحديد ميزات محددة لمجموعات البيانات أمراً هاماً للنماذج التي تؤثر على النتائج التي تدفع قيمة الأعمال.

ثانياً، أن كثرة البيانات يمكن ان تكون نعمة ونقمة. غالباً ما تكون البيانات في المؤسسة مجزأة أو موجودة في عدة أماكن. غالباً ما تأتي الاستنتاجات العلمية الدقيقة من مجموع عدة مصادر مختلفة، لكن الازدواجية يمكن أن تؤدي إلى نتائج غير متوقعة وسلبية.

ثالثاً، تعد استمرارية البيانات أمراً هاماً. مع احتمالية تغير الظروف بشكل سريع، تحتاج نماذج الذكاء الاصطناعي إلى التغذية بالبيانات الحديثة لمواصلة دفع قيمة الأعمال. من دون بيانات جديدة، سوف تتأثر قيمة النموذج. تعد النماذج "جديدة" تماماً مثل اليوم الذي تم فيه إدخال بياناتها الأساسية إلى الإنتاج وتحتاج إلى تحديث مستمر ببيانات

جديدة. يشير هذا إلى الطبيعة الدورية لهذه الخطوة نفسها ونموذج "البيانات-التدريب-الاستدلال" بالكامل. من أجل الحفاظ على الصلة، يجب وضع خطة لتحديث مجموعات البيانات الأساسية بشكل مستمر لنماذج التدريب.

أخيراً، يتطلب نقل كميات هائلة من البيانات عبر مركز بيانات بنية تحتية حوسبية فريدة. من دون البنية التحتية المناسبة لهذه المهمة، فإن تدفق عمل الذكاء الاصطناعي الخاص بك يعاني من تباطؤ في الإنتاجية والأداء قبل أن تصل حتى إلى الخطوة الأكثر طلباً - وهي التدريب.

النتيجة النهائية؟ تعتبر البيانات الواضحة وذات الصلة والحديثة مفتاحاً لإطلاق الأفكار القيمة. كل هذا العمل يأتي قبل أن تقوم بتدوير أول أعباء التدريب.



المرحلة 2: التدريب



التدريب هو مرحلة تدفق عمل الذكاء الاصطناعي التي يفكر فيها معظم العلماء غير المعنيين بالبيانات عندما يسمعون عن أعباء عمل الذكاء الاصطناعي. هذه ليست وجهة نظر غير دقيقة تمامًا. التدريب هو حيث يحدث سحر الذكاء الاصطناعي - حيث تصبح البيانات نماذج الذكاء الاصطناعي.

من دون الغوص في عمق النظرية الكامنة وراء التعلم الآلي والتعلم العميق والذكاء الاصطناعي، يمكن تلخيص التدريب كعملية تكرارية تستخدم فيها البيانات من الخطوة السابقة لإنشاء نماذج. وتُصدر هذه النماذج تنبؤات مستقبلية حول بيانات مماثلة في العالم الحقيقي. في السنوات العشر الأخيرة فقط، كان من الممكن حل المشكلات بهذه الطريقة، وذلك بفضل ظهور وحدات معالجة الرسومات أو ما يشار إليه اختصارًا GPU.

تعتبر الخوادم التي تستفيد من وحدات المعالجة المركزية (CPU) ووحدات معالجة الرسومات (GPU) الجديدة هذه "معجلة". نظرًا لأن مراكز البيانات التي تتمحور حول وحدة المعالجة المركزية (CPU) تقليديًا تتحمل المزيد من أعباء عمل الذكاء الاصطناعي، فمن الضروري تعزيز مركز البيانات باستخدام خوادم تستند إلى المعجل. تحمل أيضًا هذه الحوسبية الأكثر قوة تكلفة موارد أعلى. لهذا السبب يعد تخصيص هذه الموارد باهظة الثمن أمرًا بالغ الأهمية. التخصيص غير الصحيح يمكن أن يُفسد بسرعة مشروع الذكاء الاصطناعي.

”التدريب هو حيث يحدث سحر الذكاء الاصطناعي - حيث تصبح البيانات نماذج الذكاء الاصطناعي.“

الحاجة إلى السرعة (والدقة)

حتى في ظل الظروف المثالية، قد يستغرق التدريب على نموذج واحد فقط أيامًا أو أسابيع أو حتى أشهر. بالإضافة إلى ذلك، يتم تدريب النماذج المتوسطة عادةً من خمس إلى ست مرات قبل نشرها إلى الإنتاج. يعد تسريع الأداء للتدريب نموذجًا أمرًا لا يقدر بثمن، ولكنك تحتاج إلى مزيج من السرعة والدقة لزيادة القيمة الرئيسية لوقت التسويق.

إحدى المهام الأكثر استهلاكًا للوقت في تدفق التدريب هي إعداد وإعادة ضبط المعلمات الفائقة للنماذج. المعلمات الفائقة هي القيم التي يختارها عالم البيانات للنموذج قبل بدء التدريب. النماذج الحديثة يمكن أن تحتوي على المئات منهم. قد تستغرق عملية إعادة الضبط وإعداد التكرار ساعات، حتى مع تشغيل النموذج باستخدام عينة من مجموعات البيانات. يمكن لأتمتة عمليات البحث ذات المعلمات الفائقة، وإجراء عمليات البحث هذه بشكل متوازٍ، أن توفر لعلماء البيانات لديك أسابيع أو حتى أشهر - وتقلص الوقت اللازم لظهور نتيجة والوقت اللازم لضمان الدقة.

”يعد تسريع الأداء للتدريب نموذج لا يقدر بثمن، ولكنك تحتاج إلى مزيج من السرعة والدقة لزيادة القيمة الرئيسية لوقت التسويق.“

هل كنت تعلم؟

قد تضيع الأيام أو الأسابيع أو الأشهر إذا كان عالم البيانات في النموذج غير قادر على تحديد النجاح المبكر لمعلمت النموذج.

تتيح أدوات مثل التصور التدريبي لعلماء البيانات رؤية تقدم التدريب وتقديم تنبيهات بأن التدريب لا يتقارب. يمكن لفرق البيانات إيقاف المهمة وإعادة ضبط المعلمات وإعادة تشغيل المهمة خلال الساعات القليلة الأولى من التدريب، بدلاً من الانتظار حتى النهاية لرؤية النتيجة السيئة.



صراع قديم

تدعم المشاركة العادلة والجدولة المستندة إلى الأولويات المشاركة الديناميكية لموارد وحدة معالجة الرسومات (GPU) بين وظائف التدريب المتعددة. كما أنه يتيح استباق واسترجاع وحدات معالجة الرسومات (GPU) دون إيقاف أي وظائف. ويبقي هذا علوم البيانات باهظة الثمن منتجة؛ ولن يتم حظر أي شخص أو حرمانه من مورد وحدة معالجة الرسومات (GPU). بدلاً من ذلك، يمكنك الاستمرار في زيادة الاستفادة من هذه الموارد المكلفة إلى الحد الأقصى نظرًا لأن لديها مرونة في تبديل وحدات معالجة الرسومات (GPU) من عبء عمل إلى عبء عمل.

إن إدراك الطبيعة العالية القيمة لموارد التدريب والجهد الضروري لضمان النجاح في هذه المرحلة سوف يسفر عن النتيجة النهائية: نموذج يستمد قيمة للعمل. ومع ذلك، لا تتم المهمة بمجرد اكتمال النموذج. الآن يجب نشرها.

على عكس الكود التقليدي، تتجرف نماذج الذكاء الاصطناعي من البيانات الأساسية بمرور الوقت إذا لم يتم إعادة تدريبهم على البيانات الحديثة. لذلك، يجب إعادة تدريب أي نموذج موجود باستمرار من أجل الحفاظ على ملاءمته وفائدته. ومع ذلك، يجب أن تكون قادرًا أيضًا على طرح النماذج الجديدة وترويجها بسرعة في الإنتاج.

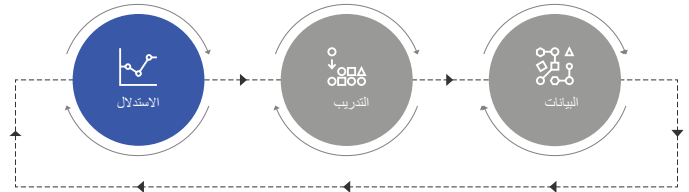
هذا يضع قادة تكنولوجيا المعلومات في قلب الصراع القديم حول تخصيص الموارد. بعض أعباء العمل أو المستأجرين أكثر أهمية من غيرهم. أو، خاضعين لاتفاقيات مستوى الخدمة (SLA) الأكثر قوة. يجب أن ينعكس هذا في طبقة جدولة المورد. يساعدك برنامج الجدولة الأسرع والأكثر فاعلية في تحقيق مستويات دقة الأعمال ذات الصلة بسرعة أكبر.

من دونها، ستعم الفوضى. يقتصر كل عالم بيانات أو مشروع على صندوق، مما يؤدي في نهاية المطاف إلى مستودع حوسبة. ويحد هذا المستخدم لجهاز واحد لا يستطيع المرء تجاوزه.

كما أنه يهدر الموارد عندما لا يستخدم هذا المستخدم الوحيد الجهاز.



المرحلة 3: الاستدلال



في الذكاء الاصطناعي، بعد النشر إلى الإنتاج هو المرحلة التي يمكنك من خلالها استخلاص رؤى من النموذج الخاص بك. هذا ما يعرف باسم الاستدلال. (البعض يسميها أيضاً التدوين). هذا هو المكان الذي تظهر فيه قيمة التعلم العميق وحيث يمكن دراسة مفاهيم الذكاء الاصطناعي الأكثر تقدماً، مثل مقاييس التفسير والإنصاف.

مرحلة الاستدلال هي في الحقيقة مجموع كل الأجزاء السابقة. إذا كانت بياناتك سيئة أو كان التدريب غير دقيق، فسيؤثر الاستدلال دون الاستدلال الصحيح، جميع الجهود السابقة لا تساوي شيء.

التحدي الأساسي في هذه المرحلة يختلف عن مرحلة التدريب. على الرغم من أن التدريب يمكن أن يتم على العديد من الدورات، ويستغرق أياماً أو أسابيع من وقت المشروع، فإن الاستدلال غالباً ما يكون عملية ثانوية فرعية تتطلب رؤية سريعة ودقيقة.

”مرحلة الاستدلال هي في الحقيقة مجموع كل الأجزاء السابقة. دون الاستدلال الصحيح، جميع الجهود السابقة لا تساوي شيء.“

مثال على الاستدلال في العمل

النظر في معالج الدفع الذي قام بتدريب نموذج للكشف عن الاحتيال الذي يحدث في معاملات المستهلك على المنصة. لن يتحمل العملاء أي تأخير غير مبرر في معالجة مدفوعاتهم. لذلك، يجب أن يضع المعالج اتفاقية مستوى الخدمة (SLA) الثانوية الفرعية وفقًا لرؤية نموذج الذكاء الاصطناعي لضمان تقدم كل معاملة بسلاسة وألا تتأثر تجربة العميل بالتحقق الاحتيالي.

للتغلب على هذه التحديات، يجب أن تكون الأجهزة الأساسية مختلفة أيضًا. بينما يتم التدريب بشكل مركزي في مراكز البيانات، يتم إجراء الاستدلال بشكل شائع على الحافة أو على أجهزة مثل الهواتف الذكية أو بالقرب من الحافة.

مثال على سيناريو قريب الحافة هو حامل خادم صغير يعمل في متجر للبيع بالتجزئة. وبشكل أكثر تحديدًا، يقوم الخادم بإنشاء رؤية في الوقت الفعلي بشأن معاملات العملاء أو تغذيات الفيديو في المتجر.

من خلال جدولة الموارد المناسبة وتخصيص الموارد، يمكنك بسرعة الحصول على رؤية ثاقبة من النماذج. من خلال استراتيجية التوسع السلسلة، يمكنك بسرعة زيادة احتياجات الاستدلال في "البرمجيات المحلية المثبتة بجهاز العميل" أو على الحافة لتلبية الطلب. على غرار مرحلة التدريب، يمكن أن تساعد القدرة على نقل مهام الاستدلال بشكل مرن في تجميع موارد مشترك في تلبية اتفاقيات مستوى الخدمة (SLA) القوية.

كما يتضح من سهم الحلقة في إطار عمل "البيانات-التدريب-الاستدلال" (DTI)، يتم تغذية البيانات التي يتم جمعها في العالم الحقيقي من خلال الاستدلال في تدفق العمل في مرحلة البيانات. يعمل إجراء الحلقة هذا على تحسين دقة النموذج باستمرار لأنه يتم تطبيق البيانات الأساسية الأعمق والأحدث. وبالتالي، تبدأ الدورة مرة أخرى.



جمع النماذج معًا

يتم توفير كل هذه المزايا بالإضافة إلى الأمان المحقق لأنظمة الطاقة، والذي يدمج بسلاسة أطر العمل مفتوحة المصدر المضمونة من قبل IBM.

إن فكرة "الذكاء الاصطناعي يفعل الصواب" تنتج أرباح هائلة. الآن بعد أن تعرفت إلى ما يلزم لتخطيط دورة الذكاء الاصطناعي ناجحة باستمرار، يبقى سؤال واحد فقط:

هل أنت مستعد للبدء؟

في كل مرحلة من مراحل تدفق عمل الذكاء الاصطناعي، من الضروري أن يكون هناك مزيج من الأفراد المناسبين والعمليات والبنية التحتية (كل من الأجهزة والبرامج) لتكون ناجحة. هذه هي المكونات الأساسية لبناء أساس قوي - المفتاح لنشر الذكاء الاصطناعي عبر عملك بالكامل.

لدى IBM® بنية تحتية للذكاء الاصطناعي تتكيف مع أولويات العمل المتغيرة، بحيث تحقق مؤسستك أهدافها في كل مرحلة من مراحل رحلة الذكاء الاصطناعي. توفر خدمة IBM Power Systems بنية تحتية للذكاء الاصطناعي رائدة في هذا المجال ومبنية خصيصًا للتعلم الآلي والتعلم العميق والاستدلال. بفضل ذلك، يمكنك:

- تغذية التفكير والقدرات الجديدة في جميع أنحاء مؤسستك.
- زيادة الثقة في القرارات التجارية على نطاق واسع.
- تحقيق أقصى استفادة من الأفراد والعمليات والمعالجات، من خلال حل مصمم لينمو مع مؤسستك.
- العثور على نتائج هامة بشكل أسرع من خلال أعلى إنتاجية للبيانات في هذا المجال وبحوث IBM، مما يجعلك في طليعة أحدث تقنيات الذكاء الاصطناعي.

تفضل بزيارة:

ibm.biz/EnterpriseAI

© حقوق الطبع والنشر لعام 2019 محفوظة لصالح شركة IBM

أنتج في الولايات المتحدة الأمريكية

يوليو 2019

إن IBM وشعار IBM و ibm.com هي علامات تجارية لشركة International Business Machines Corp المسجلة في العديد من الولايات القضائية في جميع أنحاء العالم. وقد تكون أسماء منتجات وخدمات أخرى علامات تجارية لشركة IBM أو شركات أخرى. كما تتوفر قائمة حالية بعلامات IBM التجارية على الإنترنت في قسم "معلومات حول حقوق الطبع والنشر والعلامات التجارية" على www.ibm.com/legal/copytrade.shtml

74027174-ARAE-00