ديب مايند: إلى أي مدى سيستطيع البشر تعليم الذكاء الاصطناعي؟

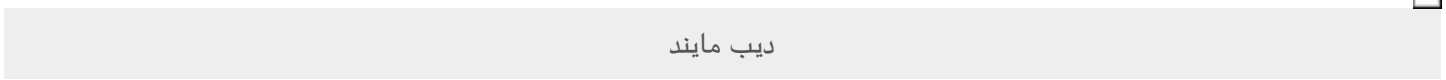


البشر يعلّمون الروبوتات

يمتلك الذكاء الاصطناعي القدرة على تطوير الإنسانية والحضارة أكثر من أي تقنية سابقة، ومع ذلك، فهو ينطوي على مخاطر ومسؤوليات كبيرة. تعمل شركة ديب مايند المملوكة من قِبل شركة ألفابيت **Alphabet** (الشركة الأم لشركة غوغل) وشركة أوبن إيه آي **OpenAI** (وهي شركة غير ربحية مختصة بأبحاث الذكاء الاصطناعي)، معاً على التخفيف من بعض هذه المخاوف، وتتعاون الشركتان مع الناس (الذين لا يملكون بالضرورة أي مهارات تقنية خاصة) لتوظيف استجاباتهم **feedback** في تعليم الذكاء الاصطناعي؛ إذ إنّ الاستجابات البشرية تساعد الذكاء الاصطناعي على أن يكون أكثر فاعلية، وتوفر أماناً وتحكماً تقنياً مُحسّناً.

ومن بين الاستنتاجات الأولى لهذا التعاون أنّ الذكاء الاصطناعي يَتعلم عن طريق التجربة والخطأ، ولا يحتاج إلى البشر لإعطائه هدفاً

نهائياً. هذا أمر جيّد، لأننا نعلم بالفعل أن ضبط هدف غير واضح يمكن أن يؤدي إلى نتائج كارثية. في التجربة العمليّة، يَستخدم النظام الاستجابات البشرية لتعليم الروبوت كيفية أداء الشقلبة.



إنّ هذا النظام غير عادي؛ لأنه أثناء اكتشاف البيئة المحيطة، يكتسب المعرفة عن طريق تدريب ما يُسمى "متنبئ المكافأة" **reward predictor** ـوهو عميل من شبكة عصبيةـ بدلاً من جمع المكافآت. ويستمر عميل تعزيز التعلم **reinforcement learning agent** باستكشاف البيئة المحيطة، ولكن الفرق هو إرسال مقاطع فيديو من سلوكه إلى شخص ما بشكل دوري، ويختار هذا الشخص السلوكَ الأفضل اعتماداً على الهدف النهائي.

تساعد هذه الاختيارات البشرية على تدريب متنبئ المكافأة، الذي يقوم بدوره بتدريب عميل التعلم. وفي النهاية، يتعلم عميل التعلم كيفية تحسين سلوكه بما فيه الكفاية لتحقيق أقصى قدر من المكافآت؛ والتي لا يمكن الحصول عليها إلا عن طريق إرضاء الإنسان.

أنظمة الذكاء الاصطناعي أكثر أماناً

يمكّن هذا النهج البشر من كشف وتصحيح أي سلوكيات غير مرغوبة، على نحو يضمن السلامة دون أن يكون مرهقاً للمشرفين. وهذا أمر جيّد، لأنهم بحاجة إلى مراجعة حوالي 0.1٪ من سلوك العميل لتعليمه. قد لا تبدو هذه النسبة مرتفعة في البداية، ولكن هذا يعني مراجعة آلاف مقاطع الفيديو، الأمر الذي يعمل الباحثون لإيجاد حل له.

يمكن أن تساعد الاستجابات البشرية الذكاء الاصطناعي أيضاً على تحقيق نتائج خارقة؛ على الأقل في بعض ألعاب الفيديو. ويقوم الباحثون الآن بتحليل سبب نجاح نظام "الاستفادة من الاستجابات البشرية" **human feedback system** وتحقيقه نتائج مذهلة في بعض المهام، ونتائج متوسطة أو حتى غير فعّالة في مهام أخرى. فعلى سبيل المثال، لا تستطيع الاستجابات البشرية مساعدة النظام في الفوز في لعبة بريك آوت **Breakout** أو كيو بيرت **Qbert**.

كما يعمل الباحثون أيضاً على حل مشكلة اختراق نظام المكافأة، حيث يؤدي التوقف المبكر للاستجابات البشرية إلى اتخاذ النظام قرارات خاطئة.

يُعد فهم هذه المشاكل أمراً ضرورياً لبناء أنظمة ذكاء اصطناعي تعمل بأمان وفاعلية كما نرغب، كما ويمكن أن تشمل الأهداف المستقبلية الأخرى تقليل كمية الاستجابات البشرية المطلوبة، أو تغيير طريقة تقديمها؛ وربما في نهاية المطاف يمكن تبادل الاستجابات "وجهاً لوجه" مما يوفّر للذكاء الاصطناعي المزيد من الفرص للتعلّم من السلوك البشري الحقيقي.

- التاريخ: 2017-08-01
- التصنيف: تكنولوجيا

#الروبوتات #الذكاء الاصطناعي

المصادر

- futurism

المساهمون

- ترجمة
  - رؤى سلامة
- مُراجعة
  - دانا أسعد
- تحرير
  - حسن شوفان
  - عماد الدين الدمري
- تصميم
  - رنيم ديب
- نشر
  - مي الشاهد