

الذكاء الاصطناعي: كيف تقلق بحكمة بشأن الذكاء الاصطناعي

يشير التقدم السريع في الذكاء الاصطناعي الخوف وكذلك الإثارة. كيف يجب أن تكون قلقاً؟

- هل يجب أن نؤتمت جميع الوظائف، حتى تلك التي نستمتع بها؟
- وهل أتممتها متعة ورفاهية للإنسان؟
- هل يجب أن نطور عقولاً غير بشرية قد تتفوق علينا عددياً، وتتفوق في الذكاء ... وقد تحل محلنا؟
- هل يجب أن نجازف بفقدان السيطرة على حضارتنا؟ "

طرحت هذه الأسئلة الشهر الماضي في رسالة مفتوحة من معهد مستقبل الحياة Future of Life Institute، وهي منظمة غير حكومية. دعا إلى التوقف لستة أشهر عن إنشاء أكثر أشكال الذكاء الاصطناعي تقدماً (AI)، ووقعها شخصيات بارزة في مجال التقنية، بما في ذلك إيلون مسك. إننا نعيش المرحلة الأبرز حتى الآن في سرعة التقدم في الذكاء الاصطناعي وإثارة القلق بشأن الأخطار المحتملة للتقنية.

فاجأت "نماذج اللغة الكبيرة" الجديدة (LLMs) - التقنية التي يقوم عليها ChatGPT - حتى المبدعين الذين طوروها، بقدراتها وردودها غير المتوقعة وتفاعلها. تتضمن هذه القدرات "الناشئة" كل شيء بدءاً من حل الألغاز المنطقية وكتابة البرامج الحاسوبية إلى التعرف على الأفلام من ملخصات الحبكة المكتوبة بالرموز التعبيرية Emojis.

ستعمل هذه النماذج على تغيير علاقة البشر بالحواسيب والمعرفة وحتى مع أنفسهم. يجادل مؤيدو الذكاء الاصطناعي بشأن قدرته على حل المشكلات الكبيرة مثل تطوير عقاقير جديدة، أو تصميم مواد جديدة للمساعدة في مكافحة تغير المناخ، أو حل تعقيدات الفيزياء وقوة اندماج المواد. بالنسبة للآخرين، فإن حقيقة أن قدرات الذكاء الاصطناعي تتفوق بالفعل على فهم مبتكريها تخاطر بإحياء سيناريو كارثة الخيال العلمي للآلة التي تتفوق على مخترعها، وغالباً ما يكون لذلك عواقب وتبعات وخيمة.

هذا المزيج المتضخم من الإثارة والخوف يجعل من الصعب الموازنة بين الفرص والمخاطر. ولكن يمكن تعلم الدروس من الصناعات الأخرى ومن التحولات التقنية السابقة. إذا

- ما الذي تغير لجعل الذكاء الاصطناعي أكثر قدرة؟
- وإلى أي درجة يجب أن نكون خائفين؟
- وماذا يجب أن تفعل الحكومات؟

في هذا الملف المخصص، نستكشف طريقة عمل نماذج اللغة الكبيرة، واتجاهها المستقبلي. اعتمدت الموجة الأولى من أنظمة الذكاء الاصطناعي الحديثة، التي ظهرت قبل عقد من الزمن، على بيانات التدريب المصنفة بعناية. بمجرد تعرضهم لعدد كافٍ من الأمثلة المصنفة، يمكنهم تعلم القيام بأشياء مثل التعرف على الصور أو نسخ الكلام. لا تتطلب أنظمة اليوم تصنيفاً مسبقاً، ونتيجة لذلك يمكن تدريبها باستخدام مجموعات بيانات أكبر بكثير مأخوذة من مصادر عبر الإنترنت. يمكن تدريب نماذج اللغة الكبيرة، في الواقع، على الإنترنت بالكامل - وهو ما يفسر قدراتهم، الجيدة والسيئة.

أصبحت هذه القدرات واضحة للجمهور على نطاق أوسع عندما تم إصدار ChatGPT في نوفمبر. استخدمه مليون شخص في غضون أسبوع؛ 100 مليون في غضون شهرين. وسرعان ما تم استخدامه لتوليد المقالات المدرسية وخطب حفلات الزفاف. دفعت شعبية ChatGPT وخطوة مايكروسوفت لدمجه في Bing، محرك البحث الخاص بها، الشركات المنافسة إلى إطلاق برامج الدردشة الآلية الذكية المنافسة أيضاً.

أسفرت بعض هذه النتائج عن ردود غريبة. اقترح Bing Chat على أحد الصحفيين أن يترك زوجته. واتهمت التقنية بالتشهير من قبل أستاذ القانون. تقدم نماذج اللغة الكبيرة إجابات تبدو للوهلة الأولى أنها علمية وذات مصداقية، ولكنها غالباً ما تحتوي على أخطاء كبيرة أو افتراءات. ومع ذلك، بدأت مايكروسوفت وجوجل وشركات التقنية الأخرى في دمج نماذج اللغة الكبيرة في منتجاتها، لمساعدة المستخدمين على إنشاء المستندات وأداء مهام أخرى.

أثار التسارع الأخير في كل من قدرات أنظمة الذكاء الاصطناعي، والوعي المتزايد بقدراتها وعيوبها، مخاوف من أن التقنية تتقدم الآن بسرعة بحيث لا يمكن التحكم فيها بأمان. ومن هنا جاءت الدعوة إلى وقفة، والقلق المتزايد من أن الذكاء الاصطناعي يمكن أن يهدد ليس فقط الوظائف والدقة العلمية والسمة، ولكن وجود البشرية نفسها.

الانقراض أو التمرد؟

يعود الخوف من سرقة الآلات للوظائف إلى قرون. ولكن حتى الآن خلقت التقنية الجديدة وظائف جديدة لتحل محل تلك التي دمرتها. تميل الآلات إلى أن تكون قادرة على أداء بعض المهام، دون غيرها، مما يؤدي إلى زيادة الطلب على الأشخاص الذين يمكنهم القيام بالمهام التي لا تستطيع الآلات القيام بها.

هل يمكن أن تكون هذه المرة مختلفة؟ لا يمكن استبعاد حدوث خلل مفاجئ في أسواق العمل، حتى لو لم يكن هناك أي مؤشر على حدوثه حتى الآن، تميل التقنيات السابقة إلى استبدال المهام التي لا تحتاج لمهارة عالية، ولكن يمكن أن تقوم نماذج اللغة الكبيرة ببعض مهام ذوي الياقات البيضاء، مثل تلخيص المستندات وكتابة التعليمات البرمجية.

كانت درجة الخطر الوجودي الذي يشكله الذكاء الاصطناعي موضع نقاش ساخن. الخبراء منقسمون. في دراسة استقصائية للباحثين في الذكاء الاصطناعي أجريت في عام 2022، اعتقد 48% أن هناك فرصة بنسبة 10% على الأقل أن تأثير الذكاء الاصطناعي سيكون "سيئاً للغاية (على سبيل المثال، الانقراض البشري)". لكن 25% قالوا إن الخطر كان 0%. قدر Median الخطر عند 5%. الكابوس هو أن الذكاء الاصطناعي المتقدم يسبب ضرراً على نطاق واسع، من خلال صنع السموم أو الفيروسات، أو إقناع البشر بارتكاب أعمال إرهابية. لا يحتاج الأمر إلى نية شريرة: يخشى الباحثون من أن يكون للذكاء الاصطناعي في المستقبل أهداف لا تتماشى مع أهداف طوروه من البشر.

لا ينبغي استبعاد مثل هذه السيناريوهات، وإن كانت جميعها تنطوي على قدر هائل من التخمين. ويتخيل الكثيرون أن الذكاء الاصطناعي في المستقبل ستوفر له الطاقة والمال وقوة الحوسبة التي سيحتاجها للوصول لهذا الأحلام والطموحات، ولكنها غير متوفرة وتمثل قيود حقيقية اليوم، وأيضاً أنه سيمكننا التحكم فيها وحرمان الذكاء الاصطناعي منها إذا خرج عما طور لأجله. علاوة على ذلك، يميل الخبراء إلى المبالغة في تقدير المخاطر في مجال تخصصهم، مقارنة بالتخصصات الأخرى. يبدو أن فرض تنظيمات صارمة، أو في "وقفة"، اليوم هو رد فعل مبالغ فيه. كما أن الوقف المؤقت لن يكون قابلاً للتنفيذ.

هناك حاجة إلى التنظيم، ولكن لأسباب أكثر منطقية من إنقاذ البشرية. تشير أنظمة الذكاء الاصطناعي الحالية مخاوف حقيقية بشأن التحيز والخصوصية وحقوق الملكية الفكرية. مع تقدم التقنية، يمكن أن تظهر مشاكل أخرى أكثر تحديداً ووضوحاً. المفتاح هو الموازنة بين الوعد بالذكاء الاصطناعي وتقييم المخاطر والاستعداد للتكيف.

تتبنى الحكومات حتى الآن ثلاث مقاربات مختلفة:

1. في أحد طرفي الطيف توجد بريطانيا، التي اقترحت نهجاً "خفيفاً" بدون قواعد أو هيئات تنظيمية جديدة، لكنها تطبق اللوائح الحالية على أنظمة الذكاء الاصطناعي. الهدف هو تعزيز الاستثمار وتحويل بريطانيا إلى "قوة عظمى في الذكاء الاصطناعي". أتبعته أمريكا نهجاً مشابهاً، على الرغم من أن إدارة الرئيس الأمريكي بايدن تبحث الآن عن وجهات نظر عامة حول الشكل الذي قد يبدو عليه دليل التنظيم والقواعد.
2. يتخذ الاتحاد الأوروبي خطأً أكثر صرامة. يصنف قانونها المقترح الاستخدامات المختلفة للذكاء الاصطناعي حسب درجة المخاطرة، ويتطلب مراقبة وكشف أكثر صرامة مع ارتفاع درجة المخاطر، على سبيل المثال، من التوصية الموسيقية إلى السيارات ذاتية القيادة. بعض استخدامات الذكاء الاصطناعي محظورة تماماً، مثل الإعلانات اللاشعورية والقياسات الحيوية عن بُعد. سيتم تغريم الشركات التي تنتهك القواعد. بالنسبة لبعض النقاد، هذه اللوائح خانقة للغاية.
3. لكن آخرون يقولون إن هناك حاجة إلى نهج أكثر صرامة. يجب على الحكومات التعامل مع الذكاء الاصطناعي، مثل التعامل مع فسخ الأدوية، مع وجود منظم مخصص، واختبار صارم وموافقة مسبقة قبل الإصدار العام. تقوم الصين ببعض من هذا، حيث تطلب من الشركات تسجيل منتجات الذكاء الاصطناعي والخضوع لمراجعة أمنية قبل إصدارها. لكن السلامة قد تكون أقل دافعاً من السياسة: أحد المتطلبات الأساسية هو أن يعكس ناتج منظمة التضامن الإسلامي "القيمة الأساسية للاشتراكية"

ماذا يجب القيام به؟

من غير المرجح أن يكون نهج اللمسة الخفيفة كافياً. إذا كان الذكاء الاصطناعي لا يقل أهمية عن التقنيات مثل السيارات والطائرات والأدوية - وهناك سبب وجيه للاعتقاد بأنها كذلك - لذا فهو مثلهم، سوف يحتاج إلى قواعد تنظيمية جديدة. وبناءً على ذلك، فإن نموذج الاتحاد الأوروبي هو الأقرب، على الرغم من أن نظام التصنيف الخاص به مُجهَد للغاية وأن النهج القائم على المبادئ سيكون أكثر مرونة. إن الإفصاح المقنع عن كيفية تدريب الأنظمة، وكيفية عملها وكيفية مراقبتها، والتي تتطلب عمليات تفتيش، سيكون قابلاً للمقارنة مع القواعد المماثلة في الصناعات الأخرى.

قد يسمح هذا بتشديد التنظيم بمرور الوقت، إذا لزم الأمر. قد يبدو من المناسب عندئذٍ وجود منظم مخصص؛ وكذلك يمكن للمعاهدات الحكومية الدولية، المشابهة لتلك التي تحكم الأسلحة النووية، أن تظهر أدلة معقولة على وجود خطر وجودي.

لرصد هذا الخطر، يمكن للحكومات تشكيل هيئة على غرار CERN، مختبر فيزياء الجسيمات، يمكن أن يدرس أيضاً السلامة والأخلاق - المجالات التي تفتقر فيها الشركات إلى الحوافز للاستثمار بقدر ما قد يرغب المجتمع.

تشكل هذه التقنية القوية مخاطر جديدة، ولكنها توفر أيضاً فرصاً غير عادية. تحقيق التوازن بين الاثنين يعني السير بحذر. يمكن أن يوفر النهج المقاس اليوم الأسس التي يمكن إضافة المزيد من القواعد عليها في المستقبل. ولكن حان الوقت للبدء في بناء تلك الأسس.