⟹ جميع الرسائل المفتوحة

أوقفوا تجارب الذكاء الاصطناعي العملاقة: رسالة مفتوحة

ندعو كافة المختبرات المعنية بالذكاء الاصطناعي إلى التوقف فورا عن تدريب أنظمة الذكاء الاصطناعي الأقوى من GPT-4

التواقيع:
31810

إضافة التوقيع

تاريخ النشر: 22 مارس/آذار 2023

يمكن أن تشكّل أنظمة الذكاء الاصطناعي ذات القدرات المنافسة للبشر مخاطر جسيمة على المجتمع والانسانية، كما يتضح من خلال عديد البحوث [1] وكما تقر به كبريات مختبرات الذكاء الاصطناعي [2]. وكما هو مذكور في مبادئ الذكاء الاصطناعي المقترنة بمؤتمر أسيلومار Asilomar AI Principles والمُعتمَدة على نطاق واسع، *فإن الذكاء الاصطناعي المتقدم قد يحدث تغييرا عميقا في تاريخ الحياة على الأرض، وعليه، ينبغي التخطيط له وإدارته بالعناية والموارد المناسبة.* ولسوء الحظ، فإن هذا المستوى من التخطيط والإدارة لم يتحقّق، على الرغم من أن الأشهر الأخيرة شهدت دخول مختبرات الذكاء الاصطناعي سباقًا غير خاضع للسيطرة يهدف إلى تطوير ونشر عقول رقمية أكثر قوة من أي وقت مضى لا يقدر أي كان – ولا حتى مبتكروها – على فهمها أو التنبؤ بمسارها أو التحكم فيها بشكل موثوق.

لقد أضحت أنظمة الذكاء الاصطناعي المعاصرة قادرة على منافسة البشر في المهام العامة [3]، وحريّ بنا أن نتساءل: هل يجب أن نترك قنوات معلوماتنا عرضة لسيل من الدعاية والأكاذيب الصادرة عن الآلات؟ هل من الحكمة أن ندع الآلة تستحوذ على كافة الوظائف، بما فيها تلك التي تعطينا إحساسا بالرّضا؟ هل يجب أن نطور عقولًا غير بشرية مآلها أن تتفوق علينا عددا وذكاء وتُنهي صلاحياتنا وتحل محلّنا؟ هل يجب أن نجازف بفقدان السيطرة على حضارتنا؟ لا ينبغي تفويض هذه القرارات لقادة تكنولوجيا غير منتخبين ولا يحسن تطوير أنظمة ذكاء اصطناعي قوية الا عندما نكون واثقين من أن تأثيراتها سوف تكون إيجابية وأن مخاطرها خاضعةٌ للسيطرة. يجب أن تكون هذه الثقة مبررة بشكل جيد وأن تتعاظم بقدر حجم تأثيرات النظام المحتملة. ينص البيان الأخير لشركة أوبن أي آي OpenAI حول الذكاء العام الاصطناعي على ما يلي: *"قد يكون من المهم الحصول في مرحلة ما على مراجعة مستقلة تسبق الشروع في تدريب أنظمة مستقبلية كما أنه من المهم أن يتوصل أصحاب الجهود الأكثر تقدما إلى اتفاق يتم بموجبه الحد من معدل نمو الحوسبة الهادفة لابتكار نماذج جديدة."* نحن موافقون على هذا ونرى أننا بلغنا هذه المرحلة.

لذلك، فإننا ندعو جميع مختبرات الذكاء الاصطناعي إلى التوقف الفوري عن تدريب أنظمة ذكاء اصطناعي أقوى من جي بي تي GPT-4 لمدة لا تقل عن ستة أشهر. يجب أن يكون هذا التوقف علنيًا وقابلًا للتحقق وشاملا لكل الجهات الفاعلة. وإن تعذر سنّ مثل هذا التوقف بسرعة، يجب على الحكومات التدخل وفرض حظر مؤقت.

يجب على مختبرات الذكاء الاصطناعي والخبراء المستقلين استخدام هذا التوقف لصياغة وتنفيذ مجموعة من بروتوكولات السلامة المشتركة اللازمة لتصميم وتطوير الذكاء الاصطناعي المتقدم يدققها ويشرف عليها خبراء خارجيون ومستقلون. يجب أن تضمن هذه البروتوكولات أن الأنظمة التي تلتزم بها آمنة بما لا يدع مجالا للشك [4]. هذا لا يعني توقفًا عن تطوير الذكاء الاصطناعي بشكل عام، بل هو مجرد تراجع عن السباق الخطير نحو نماذج مجهولة غير قابلة للتنبؤ وذات أنساق متزايدة وقدرات ناشئة.

يجب إعادة توجيه بوصلة بحوث الذكاء الاصطناعي وتطويره في اتجاه جعل النظم الحالية الصلبة أكثر دقة وأمانا وقابلية للتفسير وكذلك شفافة ومتينة، ومتسقة، وجديرة بالثقة والولاء.

بالتوازي مع ذلك، يجب على مطوّري الذكاء الاصطناعي العمل مع صانعي السياسات للتسريع بشكل دراماتيكي في تطوير أنظمة حوكمة قوية للذكاء الاصطناعي. وعلى أقل تقدير، ينبغي أن تتضمن هذه التدابير ما يلي: سلطات تنظيمية جديدة ومؤهلة مكرسة للذكاء الاصطناعي، ومراقبة أنظمة الذكاء الاصطناعي عالية القدرة والمجمّعات الكبرى للقدرات الحاسوبية وتعقبها، وأنظمة المصدر والوسم المائي للمساعدة في التمييز بين ما هو حقيقي وما مصنّع وتعقب النماذج المسرّبة، ونظام بيئي متين للتدقيق والتوثيق، والمسؤولية عن الأضرار الناجمة عن الذكاء الاصطناعي، وتمويل عام قوي لبحوث السلامة التقنية للذكاء الاصطناعي، ومؤسسات ذات موارد جيدة للتعامل مع الاضطرابات الاقتصادية والسياسية الهائلة (خاصة تلك التي تهم الديمقراطية) التي سوف يسببها الذكاء الاصطناعي.

يمكن للإنسانية أن تتمتع بمستقبل زاهر مع الذكاء الاصطناعي. بعد أن نجحنا في إنشاء أنظمة ذكاء اصطناعي قوية، حقّ لنا الآن الاستمتاع بـ "صيف ذكاء اصطناعي" نحصد فيه المكافآت، ونهندس هذه الأنظمة لتحقيق منفعة واضحة للجميع، ونمنح المجتمع فرصة للتكيّف. لقد سبق وأن توقف المجتمع مؤقتا عن تطوير تقنيات أخرى كانت تحمل آثاراً كارثية محتملة عليه [5]. الآن، يمكننا القيام بنفس الشيء. دعونا نستمتع بصيف طويل للذكاء الاصطناعي، وليس الاندفاع والسقوط دون استعداد.

هذه بعض الأسئلة الشائعة إجابة على تساؤلات ونقاشات تم التطرق إليها في وسائل الإعلام وغيرها. يمكنك الاطلاع على هذه الأسئلة هنا.

بإضافة إلى هذه الرسالة المفتوحة، نشرنا مجموعة من التوصيات السياسية التي يمكن الاطلاع عليها هنا:

صنع السياسات أثناء فترة التوقف عرض الورقة 12 أبريل/نيسان 2023

يمكنك أيضا تحميل الرسالة المفتوحة في شكل بي دي أف PDF