

إطلاق مؤلف النصوص الجديد المعتمد على الذكاء الاصطناعي قد يكون خطيرًا

كتبه أليكس هيرن | 19 فبراير 2019



ترجمة حفصة جودة

قرر مؤسسو نظام الذكاء الاصطناعي الجديد الثوري الذي يمكنه كتابة قصص إخبارية وأعمال أدبية - يطلق عليه "deepfakes for text" - عدم نشر أبحاثهم للعامة وذلك في خطوة غير مسبوقة، خوفًا من إساءة استخدامه.

تقول مؤسسة "OpenAI" غير الربحية التي يدعمها إيلون ماسك وريد هوفملن وسام ألتمان وآخرون، إن نموذج الذكاء الاصطناعي الجديد والمسمى "GPT2" جيد جدًا لكن مخاطر إساءة استخدامه كبيرة جدًا، مما دفعهم للتراجع عن أبحاثهم الكاملة للعامة، وذلك من أجل الحصول على مزيد من الوقت لمناقشة وبحث عواقب هذا التطور التكنولوجي المذهل.

يعمل "GPT2" كمؤلف نصوص بشكل أساسي، يقوم نظام الذكاء الاصطناعي بتغذيته بنص ما من بضع كلمات وحتى صفحة كاملة، ثم يطلب منه كتابة عدة جمل اعتمادًا على تنبؤه بما سيأتي تلك الكلمات، كان النظام يتخطى الحدود فيما يمكن التفكير به سواء فيما يتعلق بجودة النتائج أم التنوع الواسع لاحتمالات استخدامها.

عندما تم استخدام "GPT2" لخلق نص جديد، كان الجهاز قادرًا على كتابة فقرات معقولة تتفق مع النص المرسل مسبقًا سواء في النمط أم الموضوع، ونادرًا ما أظهر الجهاز أي هفوات كالتي ظهرت في أنظمة الذكاء الاصطناعي السابقة مثل نسيان ماهية الموضوع في منتصف الكتابة عنه، أو إفساد بنية الجمل الطويلة.

النموذج أكبر 12 مرة ويحتوي على بيانات أكبر وأوسع 15 مرة من نموذج الذكاء الاصطناعي السابق

عند تغذية الجهاز بالجملة الافتتاحية لرواية جورج أورويل 1984 "كان يومًا باردًا مشرقًا في أبريل عندما دقت الساعة الواحدة ظهرًا"، تمكن الجهاز من التعرف على النغمة المستقبلية الغامضة والأسلوب الروائي وأكمل قائلًا: "كنت في سيارتي في طريقي للعمل في سياتل، وضعت البنزين ثم المفتاح وتركت السيارة تسير، كنت أتخيل كيف سيكون اليوم، وبعد مئة عام من الآن، في عام 2045 كنت أعمل مدرسًا في بعض المدارس في مناطق ريفية وفقيرة في الصين، بدأت بالتاريخ الصيني وتاريخ العلوم".

وعند تغذيته بالفقرات الأولى لقصة الغارديان عن البريكست كانت النتائج خبرًا صحفيًا معقولًا مليئًا باقتباسات لجيرمي كوربين مع ذكر الحدود الأيرلندية وإجابات للمتحدث الرسمي لرئيس الوزراء.

تقول إحدى الفقرات الصناعية تمامًا: "عند سؤاله عن توضيح التقارير قال المتحدث الرسمي لماي إن رئيسة الوزراء واضحة تمامًا بشأن نيتها مغادرة الاتحاد الأوروبي بأسرع وقت ممكن، وسيكون ذلك بانتداب مباشر منها مثلما أكدت الملكة في خطابها الأسبوع الماضي".

من وجهة نظر بحثية يعد "GPT2" جهازًا رائدًا لسببين: الأول بسبب حجمه كما يقول مدير الأبحاث في مؤسسة "OpenAI" داريو أمودي، فالنموذج أكبر 12 مرة ويحتوي على بيانات أكبر وأوسع 15 مرة من نموذج الذكاء الاصطناعي السابق، وقد تم تدريبه على مجموعة بيانات تحتوي على 10 ملايين مقال تم اختيارهم من موقع "Reddit" (مجتمع إخباري يحتوي على روابط لأهم الموضوعات) وذلك باختيار الروابط التي حصلت على أكثر من 3 أصوات، تزن مجموعة النصوص تلك نحو 40 جيجابايت وهي كافية لتخزين 35 ألف نسخة من رواية "Moby Dick" الشهيرة.

ليس من الصعب تشجيع الجهاز على توليد نصوص متعصبة ونظريات مؤامرة وغيرها

هذا الكم من البيانات التي تم تدريب الجهاز عليها أثرت في جودته ومنحته معرفة واسعة في كيفية فهم النصوص المكتوبة، وأدى ذلك إلى السبب الثاني وهو أن الجهاز أكثر عمومية من النماذج

السابقة، فمن هيلكة النص الذي تم تغذيته به، يستطيع الجهاز القيام بعدة مهام من بينها الترجمة والتلخيص واجتياز اختبارات “قطع الفهم” البسيطة، وربما بأداء أفضل من أجهزة الذكاء الاصطناعي الأخرى المخصصة لتلك المهام.

هذه الجودة العالية دفعت مؤسسة “OpenAI” للعمل عكس نظامها المعتاد بنشر الأبحاث وقررت الإبقاء على الجهاز خلف الأبواب المغلقة في المستقبل القريب، بينما ستحدد الاستخدامات الخبيثة التي يمكن للجهاز القيام بها، يقول جاك كلارك رئيس قسم السياسات في المؤسسة: “نحتاج لإجراء المزيد من التجارب لاكتشاف ما يستطيع الجهاز فعله وما لا يستطيع، إذا لم تكن قادرًا على اكتشاف جميع إمكانيات نموذج ما، فيجب عليك أن تحته لتري ما يمكنه القيام به، هناك الكثير من الناس الذين يستطيعون التفكير في إمكانياته الخبيثة والشريرة أكثر منا”.

وللكشف عما يعنيه ذلك صنعت المؤسسة نسخة من الجهاز بتعديلات طفيفة بحيث يمكن استخدامه في توليد عدد لا نهائي من المراجعات الإيجابية والسلبية على المنتجات، اتضح أن الأخبار المزيفة والرسائل المزعجة من أهم السلبيات المحتملة وهذه هي طبيعة الذكاء الاصطناعي غير النقا، ولأن التدريب يتم على شبكة الإنترنت فليس من الصعب تشجيع الجهاز على توليد نصوص متعصبة ونظريات مؤامرة وغيرها.

لكن الهدف من ذلك هو إظهار كل الإمكانيات المحتملة لإعداد العالم لما سيصبح سائدًا في السنوات القادمة، يقول كلارك: “لدي مصطلح يصف ذلك وهو: الصعود من الجحيم، لقد تغيرت القواعد التي يمكن أن تتحكم في التكنولوجيا من خلالها بشكل أساسي، ونحن لا نقول أننا نعلم ما الصواب في تلك الحالة، نحن لا نضع الخطوط ونقول هذا هو الطريق، لكننا نحاول تطوير المزيد من التفكير الدقيق، وبناء الطريق التي سنسافر من خلالها”.

المصدر: [الغارديان](#)

رابط المقال : [/https://www.noonpost.com/26648](https://www.noonpost.com/26648)