هل الذكاء الاصطناعي التوليدي يشكّل خطورة أكثر ممّا نعتقد؟

محمد معاذ

2023-05-15

خلال الأشهر الأخيرة الماضية، تمّ كتابة الكثير من المقالات والأوراق حول مخاطر الذكاء الاصطناعي التوليدي. ويمكن حصرها في ثلاثة محاور أساسية، إلا أنها قد لا تعكس الخطر الأكبر الذي قد يعترض طريقنا كبشر. وقبل الخوض في غمار الخطر الخفي المتمثّل في الذكاء الاصطناعي التوليدي.

منظمة المجتمع العلمي العربي

لا بدّ أولًا من تلخيص تلك المخاطر والتحذيرات الشائعة.

مخاطر تتعلّق بالوظائف

إنّ التطوّرات في الذكاء الاصطناعي تأتي متبوعة بمخاوف تتعلّق بالوظائف والعمل. ولا تمثّل أحدث موجة من نماذج الذكاء الاصطناعي مثل تشات جي بي تي خروجًا عن هذه القاعدة. ويمكن للذكاء الاصطناعي التوليدي الآن إنتاج أعمال على المستوى البشري تتراوح من الأعمال الفنية والمقالات إلى التقارير العلمية. وهذا سيؤثر بشكل كبيرٍ على سوق العمل، لكننا نفترض أن ذلك لن يمثل مخاطرة كبيرة، حيث تتكيّف تعريفات الوظائف مع قوة الذكاء الاصطناعي. نعم، قد يكون الأمر مؤثرًا لفترة، لكنه لا يختلف عن كيفية تكيّف الأجيال السابقة مع الاستعانة بأدواتٍ توفر المزيد من الكفاءة في العمل.

مخاطر المحتوى المزيف

بات بإمكان الذكاء الاصطناعي التوليدي إنشاء أعمالٍ فنية كما لو أنها من صنيعة البشر، بما في ذلك المقالات والمقالات والأوراق ومقاطع الفيديو المفبركة. إنّ المعلومات المضلّلة ليست مشكلة جديدة، لكن الذكاء الاصطناعي التوليدي سيسمح بإنتاجها بكمياتٍ كبيرة، وبمستويات لم يكن لها مثيل. قد يشكّل ذلك خطرًا كبيرًا، لكن نعتقد أن هناك إمكانية للتحكم به نسبيًا. حيث يمكن تحديد المحتوى المزيف، إمّا من خلال تقنية العلامات المائية

(watermarking) التي قد تحدّد محتوى الذكاء الاصطناعي المزيف، أو عبر الاستعانة بأدواتٍ مضادة قائمة على الذكاء الاصطناعي والتي يتم تدريبها بغية تحديد المحتوى الناشئ عن هذه التقنية.

مخاطر الآلات "الواعية"

يشعر العديد من الباحثين بالقلق من أنّ أنظمة الذكاء الاصطناعي سيتم توسيعها إلى مستوى أعلى حيث يتم تطوير ما يطلق عليه البعض "إرادة خاصة بها" وقد تتخذ إجراءات تتعارض مع المصالح البشرية، أو حتى تهدد الوجود البشري. قد يمثّل هذا خطرًا حقيقيًا فعليًا على المدى الطويل. ولكن مع ذلك، نفترض أن أنظمة الذكاء الاصطناعي الحالية لن تصبح "واعية" دون إجراء تحسينات هيكلية كبرى على التقنية. وطبعًا، هذا إن حصل. وبالتالي، فهذا ليس الخطر الأكثر إلحاحًا الذي يقف حائلًا في الوقت الراهن.

لكن، ما الذي يدعو للقلق أكثر بشأن ظهور الذكاء الاصطناعي التوليدي؟

نعتقد أنّ المكان الذي يخطئ فيه معظم خبراء السلامة، بما فيهم صانعي السياسات، هو أنهم ينظرون للذكاء الاصطناعي التوليدي بشكلٍ رئيسي على أنه أداة لإنشاء محتوى تقليدي على نطاقٍ واسع. في حين أن هذه التقنية تعدّ ماهرة جدًا في طرح المقالات والصور ومقاطع الفيديو. والمشكلة الأكثر أهمية هي أن الذكاء الاصطناعي التوليدي سيطلق العنان لشكلٍ جديد من الوسائط يتسّم بالتخصّص والتفاعل الكامل، وربما يكون أكثر تلاعبًا من أي شكل محتوى عرفناه حتى الآن.

عصر الوسائط التوليدية التفاعلية

إنّ الميزة الأكثر خطورة للذكاء الاصطناعي التوليدي لا تتمثّل في قدرته على نشر مقالات ومقاطع فيديو مزيفة على نطاقٍ واسع، وإنّما في إمكانية إنتاج محتوى تفاعلي ومتكيّف مع كل مستخدم بغية تحقيق أكبر قدر من التأثير الإقناعي. وفي هذا السياق، يمكن تعريف الوسائط التوليدية التفاعلية على أنها مواد ترويجية يتم إنشاؤها أو تعديلها في الوقت الحقيقي وذلك لزيادة التأثير بناءً على البيانات الشخصية حول المستخدِم.

إنّ الإعلانات التوليدية المركّزة هي استخدام الصور ومقاطع الفيديو وغيرها من أشكال المحتوى المعلوماتي التي تبدو وكأنها إعلانات تقليدية ولكنها موجّهة في الوقت الحقيقي للمستخدمين الأفراد. وسيتم إنشاء هذه الإعلانات مباشرةً بواسطة أنظمة الذكاء الاصطناعي التوليدية بناءً على أهداف التأثير إلى جانب البيانات الشخصية التي يتم الوصول إليها للمستخدم المحدّد. وقد تشمل البيانات الشخصية: العمر ، والجنس، والمستوى التعليمي، بالإضافة إلى مجالات الاهتمام، والقيم، والميول الشرائية وما إلى ذلك. وسيقوم الذكاء

الاصطناعي التوليدي بتخصيص التصميم، والصور، والرسائل الترويجية وذلك لزيادة الفعالية إلى أقصى حد على هذا المستخدم. بمعنى آخر، يمكن تخصيص كل التفاصيل في الوقت الحقيقي وذلك لزيادة التأثير الخفي على الفرد. ونظرًا لأنّ المنصات التقنية يمكنها تتبّع تفاعل المستخدم، سيكون بإمكانها تعلّم التكتيكات التي تعمل بشكلٍ أفضل على المستخدم. ومؤخرًا أعلنت كلّ من شركة ميتا و جوجل عن خططٍ لاستخدام الذكاء الاصطناعي التوليدي في إنشاء الإعلانات عبر الإنترنت وهذا ما سيفتح الباب على منافسة شديدة لتخصيص المحتوى الترويجي بأكثر الطرق فعالية الممكنة.

وكل ذلك يقود إلى تأثير إعلانات المحادثة أو التسويق التفاعلي، وهي تقنية يتم من خلالها نقل التأثير من خلال محادثاتٍ تفاعلية بدلًا من المحتوى الكتابي أو مقاطع الفيديو التقليدية. وستجري المحادثات من خلال روبوتات المحادثة مثل "تشات جي بي تي" أو "بارد" أو من خلال أنظمة قائمة على الصوت مدعومة بنماذج اللغات الكبيرة (LLMs)، حيث سيتم دمجها في مواقع الويب والتطبيقات والمساعدين الرقميين. وهنا، يمكن أن يكون المستخدم مستهدفًا في رسائل خفية منسوجة في الحوار مع الأهداف الترويجية. ويمكن للرعاة أن يضخّوا هذه الرسائل دون ملاحظتها حتى.

بالإضافة إلى ذلك، يمكن لهذه الأنظمة الذكية توثيق مدى نجاح المحادثات السابقة في إقناع المستخدم، ومعرفة التكتيكات الأكثر فعالية عليه، مثل مدى الاستجابة للنداءات المنطقية أو للحجج العاطفية وغيرها. وبالتالي، فنحن أمام أشبه بـ "حرباء رقمية" لا تقدم أي رؤى حول المخرجات الخاصة بها، ولكنها مسلحة ببيانات مكثفة حول المستخدم ورغباته وميوله الشخصية. وهنا، يكمن الخطر الحقيقي في دفع المعلومات المضللة إلى الواجهة، والتحدّث عن معتقداتٍ خاطئة أو أيديولوجيات متطرفة قد تكون مرفوضة في الحالة الطبيعية.

وبالمحصلة، فإنّ نشر أنظمة الذكاء الاصطناعي التوليدي بدون سياساتٍ واضحة، وأطر حماية ذات مغزى للمستخدمين، سيجعلهم عرضة لممارسات تتراوح بين الإكراه الخفي والتلاعب الصريح من قِبل هذه الأنظمة.

تواصل مع الكاتب: m.maaz@arsco.org

يسعدنا أن تشاركونا أرائكم وتعليقاتكم حول هذهِ المقالة عبر التعليقات المباشرة بالأسفل أو عبر وسائل التواصل الإجتماعي الخاصة بالمنظمة

src= src= src= src= src= src= src=