شركة OpenAI تعترف بالإخفاق بمشروع مكافحة الغش عبر الذكاء الاصطناعي

تخلَّت شركة "أوبن إيه آي" OpenAI، التي ابتكرت تطبيق "شات جي بي تي" ChatGPT ، عن تطوير أداتها المصممة للكشف عن النصوص المكتوبة بالذكاء الاصطناعي، في خطوةٍ وصفها خبراء بأنها نكسة لمساعي مكافحة الغش والتضليل والكشف عن السرقات العلمية والأدبية، بحسب صحيفة The Times البريطانية.

كانت شركة "أوبن إيه آي" قد أطلقت مشروع برنامجها المسمى Classifier، في يناير/كانون الثاني الماضي، وقالت إن الغاية منه مكافحة "الحملات الرقمية التي تنشر معلومات مضللة؛ والاحتيال الأكاديمي؛ والاستعانة بتطبيقات المحادثة بالذكاء الاصطناعي مكان المستخدمين البشريين"، فضلاً عن استخدامه في تحديد النصوص المكتوبة بالذكاء الاصطناعي، بعد أن أصبحت تأثيرات هذه التقنية "قضية شاغلة بين المعلمين".
لكنها ما لبثت أن ألغت المشروع بهدوء بسبب "انخفاض معدل الدقة" في البرنامج.

وزعمت الشركة حين أطلقت البرنامج أن "معدل نجاحه في تحديد النصوص المكتوبة بالذكاء الاصطناعي من

بين النصوص المكتوبة بهذه التقنية يبلغ 26%"، لكنها لم تقدم تفاصيل أخرى عن تطور البرنامج، ولعلَّها نشرت بعض التحديثات، لكن التخلي عن التطبيق يدل على أنه لم يشهد تحسناً كبيراً.

لا تقتصر عيوب هذه البرامج على انخفاض معدل نجاحها في الكشف عن نصوص الذكاء الاصطناعي، بل يزيد على ذلك أنها قد تستخدم في اتهام الطلاب ظلماً.

حيث قد أبلغ طلاب أمريكيون بالفعل عن اتهامهم زوراً بالغش بعد أن استخدمت كلياتهم موقع "تورنيتين" العلمية والسرقات الانتحال لمكافحة GPTZero "زيرو تي بي جي" وأداة Turnitin.
مع ذلك، فإن قرار التخلي عن المشروع ضربةٌ للجهود التي تبذلها المؤسسات الأكاديمية لمكافحة المعلومات المضللة، والمساعي المتواصلة للبحث عن حلولٍ تقنية لمشكلة الغش والأخبار الكاذبة.

حملة لمكافحة التضليل
في الأسبوع الماضي، تعهدت 7 شركات رائدة في مجال الذكاء الاصطناعي أمام مسؤولي البيت الأبيض بتطوير "آليات تكنولوجية قوية للتعرف على المحتوى الذي استُخدم الذكاء الاصطناعي في إنشائه، وقد تشمل هذه الآليات نظام العلامات المائية".

ونجحت كثير من المساعي المبذولة لتطوير الاعتماد على العلامات المائية في الكشف عن الصور ومقاطع الفيديو والصوت المنشأة بواسطة أدوات الذكاء الاصطناعي، أما النصوص المكتوبة بهذه الأدوات فلطالما اعترض الكشف عنها عوائق تقنية أكثر صعوبة.

وخلصت دراسة أجرتها جامعة ميريلاند الأمريكية هذا العام، إلى أن "هذه الكواشف الحديثة لا يمكنها الكشف بدرجة موثوقة عن النصوص المكتوبة بنماذج اللغة الكبيرة [أدوات الذكاء الاصطناعي] في الاستخدام العملي".

من جانبه، شكَّك سام ألتمان، رئيس شركة "أوبن إيه آي"، في فاعلية أدوات الكشف عن سرقة النصوص، وقال في يناير/كانون الثاني الماضي، إن الشخص "العازم" على الأمر لن يعدم أن يجد وسيلةً ما للتهرب من هذه الأدوات".
أضاف: "نحن نعيش في عالم جديد يتعين علينا فيه أن نتكيف مع النصوص المكتوبة بالذكاء الاصطناعي"، فعلى الرغم من أن النماذج اللغوية "وسيلة شديدة التقدم في هذا السياق"، فإنها تقدم أيضاً "مزايا غير مسبوقة".

من جهة أخرى، قال سلمان خان، المؤسس والرئيس التنفيذي لـ"أكاديمية خان" التعليمية على الإنترنت، هذا الأسبوع، إن النجاح في إنشاء أدوات فعالة للكشف عن سرقات النصوص بالذكاء الاصطناعي أمر "مستحيل"، وحثّ المدارس والجامعات على استيعاب أدوات الذكاء الاصطناعي بدلاً من التصدي لها.

وقال خان لصحيفة "التايمز" البريطانية: "معظم النصوص المكتوبة بالذكاء الاصطناعي التوليدي نصوص مستحدثة. أما أدوات الكشف عن الانتحال وسرقات النصوص فتبحث عن عبارات أو فقرات اطلعت عليها من قبل، ومن ثم فإننا لن نجد أدوات فعالة" للكشف عن النصوص.

وتابع خان: "بعض الناس يقولون إن إحدى الوسائل الممكنة أن تفرض كبرى شركات النماذج اللغوية علامات مائية على النصوص المكتوبة بها. لكن الواقع يشهد بأنه: أولاً: من السهولة بمكانٍ أن تكسر العلامة المائية؛ وثانياً: من السهولة بمكانٍ أن تبتكر أداة يمكنها التعرف على العلامة المائية ثم تُستخدم لتشويشها".