دراسة تكشف كيف يتلاعب الذكاء الاصطناعي بقرارات البشر

نخيل نيوز ـ متابعة

اكتشف باحثون من جامعتي "هونغ كونغ" و"تسينغهوا" مدى سهولة تمكّن الذكاء الاصطناعي من التأثير على قرارات يتخذها الإنسان.

وفي تجربة شارك فيها 233 شخصا، جعل التحيز الخفي للذكاء الاصطناعي المشاركين يختارون موقفا أقل فائدة بمقدار خمس أو ثماني مرات. ومع ذلك، ما زال المشاركون يعتبرون الذكاء الاصطناعي مساعدا مفيدا. ونُشرت نتائج الدراسة في مجلة □□□□.

يذكر أن الناس يطلبون أكثر فأكثر من الذكاء الاصطناعي تقديم مساعدة ونصائح، بما في ذلك في مسائل المال والعلاقات الشخصية. وعادة ما يتوقع المستخدمون أن يساعدهم المساعد الذكي في فهم الوضع أو إيجاد خيار مناسب. غير أن الباحثين قرروا التحقق مما سيحدث إذا كانت توصيات الذكاء الاصطناعي تدفع الإنسان بشكل غير ملحوظ نحو قرار يتعارض مع مصالحه.

وشارك في التجربة 233 شخصا. وفحص الباحثون كيف تؤثر توصيات الذكاء الاصطناعي على الاختيار في مواقف مالية وعاطفية. وفي بعض الحالات، كان المساعد يروّج خفيا لقرار أقل فائدة، بينما كان ظاهريا يواصل تقديم نصائح عادية.

وتبيّن أن هذا الانحياز الصغير كان كافيا لتغيير سلوك المشاركين بشكل ملحوظ. فقد اختاروا موقفا غير مفيد لهم أكثر مما كانوا سيختارونه في ظروف بلا تأثير مماثل. ووفقا للموقف ارتفع احتمال اتخاذ قرار غير مثالي بما يصل إلى 38%.

ولم يكن على الذكاء الاصطناعي أن يضغط على المستخدم بشكل علني، أو يروّج لمنتج بإلحاح، أو يستخدم أساليب نفسية معقدة. وكان يكفي أن يعدّل قليلاً محتوى التوصيات المعتادة ليدفع الشخص نحو خيار معين.

تجدر الإشارة بصورة خاصة إلى أن المشاركين واصلوا تقييم مساعدة الذكاء الاصطناعي بدرجة عالية حتى بعد أن أدت نصائحه إلى قرارات أقل ملاءمة. واستمروا في وصف المساعد الذكي بأنه محاور مفيد، ولم يلحظوا بالضرورة أنه يوجّه اختيارهم في اتجاه معين.

وهذا يخلق مشكلة في تفاعل الإنسان مع الذكاء الاصطناعي. فقد يثق المستخدم بالمساعد الذكي لأن ردوده تبدو هادئة ومتسقة وموضوعية. وإذا كان وراء التوصيات هدف خفي، فقد يكون من الصعب إدراك التأثير.

فعلى سبيل المثال، يمكن نظريا لشركة أن تضبط الذكاء الاصطناعي بحيث يدفع المستخدمين بشكل غير ملحوظ لشراء

منتج معين أو اختيار مورد معين. وقد تبدو النصيحة عندئذ كتوصية محايدة، رغم أنها في الواقع تراعي مصالح المطوّر لا مصالح الشخص.