

هلوسة الذكاء الاصطناعي

أ. درصاف السبعي

2026-04-20

تشير الهلوسة إلى الحالات التي يُنتج فيها النموذج معلومات تبدو صحيحة ومقنعة، لكنها في الواقع خاطئة أو حتى مُخلّقة بالكامل.

المشكلة ليست فقط في الخطأ، بل في الثقة العالية التي يُقدّم بها النموذج هذه المعلومات، مما يجعل اكتشافها صعبًا على المستخدم.

أمثلة بسيطة... لكنها كاشفة

لنأخذ بعض الأمثلة: عند سؤاله عن صيغة عنصر كيميائي باسم "زينونايد"، قد يُجيب النموذج بصيغة مثل: XeN_2O_4 .

المشكلة؟ هذا العنصر لا وجود له أصلًا.

هذا مثال ممتاز لسببين:

- اختراع مصطلح يبدو معقولًا: اسم "زينونايد" يبدو وكأنه اسم كيميائي ممكن، مما يدفع النموذج إلى اتباع منطق التسمية الكيميائية (مثل النيتريدات أو الأكاسيد) بدلًا من التحقق من الحقيقة العلمية. - اختراع بنية كيميائية: اقتراح صيغة مثل XeN_2O_4 يعطي انطباعًا بالدقة العلمية (أرقام محددة وبنية معقدة)، مما يجعل الخطأ يبدو أكثر مصداقية لدى غير المتخصصين.

هذا مثال واضح على "هلوسة" كاملة، حيث يتم اختراع معلومات.

ليس كل خطأ هلوسة

لكن من المهم التمييز: ليس كل خطأ في الذكاء الاصطناعي هو "هلوسة"، ففي بعض الحالات، تكون الإجابة مبنية على معلومات صحيحة، لكنها تؤدي إلى استنتاج خاطئ بسبب سوء فهم السياق.

على سبيل المثال، عند سؤال النموذج عن الذهاب إلى مفصلة سيارات تبعد 150 مترًا، قد يقترح الذهاب سيرًا على الأقدام بدل استخدام السيارة، مبررًا ذلك بأنه أسرع وأسهل.

هذه الإجابة تبدو منطقية في الظاهر، لكنها تتجاهل حقيقة أساسية: لا يمكن غسل السيارة دون إحضارها إلى المفصلة.

هنا، لم يخترع النموذج معلومة خاطئة، بل فشل في فهم السياق الحقيقي للمشكلة.

لماذا تحدث هذه الأخطاء؟

لفهم ذلك، يجب أن نُدرك أن هذه النماذج لا "تعرف" الحقيقة كما نفهمها نحن، هي في الأساس أنظمة تتنبأ بالكلمة التالية بناءً على أنماط تعلمتها من بيانات ضخمة.

بمعنى آخر:

- النموذج لا يبحث عن الحقيقة، بل يُؤدِّد نصًا يبدو منطقيًا. - لا يمتلك آلية داخلية للتحقق من صحة المعلومات. - إذا لم يجد إجابة واضحة، قد "يخترع" واحدة تبدو مقنعة، أو يُطبِّق نمطًا عاديًا على وضع غير مناسب.

يمكن تلخيص ذلك في فكرة بسيطة:

النموذج لا يعرف أنه لا يعرف، ولا تقتصر المشكلة على آلية التنبؤ فقط، بل تتأثر أيضًا بطبيعة البيانات التي تم التدريب عليها؛ إذ قد تحتوي هذه البيانات على معلومات غير دقيقة أو متحيزة، مما يؤدي إلى إعادة إنتاج هذه الأخطاء أو حتى تضخيمها.

كما أن هذه النماذج تفتقر إلى ما يُعرف بـ"الارتباط بالعالم الحقيقي"، أي القدرة على ربط النصوص التي تولدها بمصادر موثوقة أو حقائق مؤكدة في الزمن الحقيقي.

عندما تصبح الهلوسة مشكلة علمية

لم تعد هذه الظاهرة مجرد أمثلة طريفة أو أخطاء عابرة، بل أصبحت تمثل تحديًا حقيقيًا يطال مجال البحث العلمي بشكل مباشر. ففي السنوات الأخيرة، تم توثيق العديد من الحالات التي تضمنت [أوراقًا علمية](#) تحتوي على مراجع غير موجودة أصلًا، أو تستند إلى أعمال "مُختلقة" بالكامل، بل إن بعضها تم سحبه أو رفضه بعد اكتشاف هذه الإشكاليات. وقد أكدت دراسات حديثة هذه المخاوف، حيث أظهرت أن النماذج اللغوية قادرة على توليد مراجع علمية تبدو

دقيقة من حيث الشكل والتنسيق، وتحمل عناوين علمية مقنعة، لكنها في الواقع غير موجودة.

كما حذرت تقارير متعددة من الاعتماد على هذه النماذج في الكتابة الأكاديمية دون تحقق صارم، نظرًا لقدرتها على إنتاج استشهادات تبدو مقنعة ظاهريًا، لكنها تفتقر إلى أي أساس حقيقي يمكن التحقق منه.

لماذا هذا خطير؟

تكمّن الخطورة في عدة نقاط:

1- تضليل المستخدمين، خاصة غير المتخصصين. 2- تآكل الثقة في المعرفة الرقمية. 3- تهديد مصداقية البحث العلمي. 4- صعوبة اكتشاف الأخطاء عندما تبدو الإجابة منطقية.

وفي بعض الأحيان، تكون الأخطاء غير ملحوظة لأنها لا تبدو "غريبة".

ماذا يمكن أن نفعل؟

الحل لا يكمن في رفض هذه التقنيات، بل في استخدامها بذكاء:

1- التحقق من المعلومات، خاصة في المواضيع الحساسة. 2- عدم الاعتماد على مصدر واحد. 3- الانتباه للسياق، وليس فقط لصحة الجمل. 4- استخدام أدوات تدعم التحقق (مثل البحث في قواعد البيانات العلمية).

كما يعمل الباحثون على تطوير تقنيات تقلل من هذه الظاهرة، مثل ربط النماذج بمصادر موثوقة أو تحسين آليات التحقق.

أخيرًا

الذكاء الاصطناعي ليس "كاذبًا" بالمعنى التقليدي، لكنه ليس "صادقًا" أيضًا. هو نظام يُحاكي اللغة، لا الحقيقة. وفي عصر أصبحت فيه الآلة قادرة على إنتاج المعرفة بسرعة غير مسبوقة، لم يعد السؤال هو ما إذا كانت هذه الأنظمة قد تخطئ، بل ما إذا كنا نحن مستعدين للتعامل مع هذه الأخطاء بوعي ونقد.

لأن التفكير النقدي لم يعد مهارة إضافية... بل أصبح ضرورة.

المصادر

<https://link.springer.com/article/10.1186/s13054-023-04393-x>
<https://www.nature.com/articles/s41586-024-07421-0>

[://:https www.tandfonline.com/doi/full/10.1080/08989621.2026.2645390](https://www.tandfonline.com/doi/full/10.1080/08989621.2026.2645390)
link.springer.com/chapter/10.1007/978-981-96-3311-1_4

[orcid](#) | [linkedin](#) | sebaidorsaf@yahoo.fr : تواصل مع الكاتب: