جيميناي: الخطوة التالية في تطور الذكاء الاصطناعي

كتبه مصطفى أحمد | 10 ديسمبر ,2023

توقع المجتمع التكنولوجي أن ينتهي عام 2023 نهاية هادئة قليلًا بعد كمّ الأحداث التي توالت خلاله، بداية من إطلاقات نماذج الذكاء الاصطناعي، مرورًا بأزمة إيلون ماسك الأخيرة، ثم مهزلة سام آلتمان الإدارية، وتمنى رواد السيلكون فالي فترة إجازات هانئة، خاصة في قطاع الذكاء الاصطناعي الذي وصل نهاية العام إلى أقصى ما كنا نتوقع في بداياته، إلا أن جوجل (ألفابيت) كان لها رأي آخر.

لم تكتفِ جوجل خلال هذا العام بإطلاق نموذج "بارد (Bard)" وربطه بالكثير من تطبيقاتها، بالإضافة إلى إطلاق هاتف "بيكسل 8 برو" الذي أُطلق عليه "أول هاتف مدعّم بتقنيات الذكاء الاصطناعي"، ذلك الهاتف أبهرنا بإمكاناته الهائلة في معالجة النصوص والصور وحتى مقاطع الفيديو والتعديل عليها، لكنها قررت أن تفاجئ الجميع بنموذج ذكاء اصطناعي ضخم متعدد الوسائط ذي قدرات هائلة، وكان من بين ما فاجأتهم، وهم أهمهم، منافسوها في مجال الذكاء الاصطناعي، ميكروسوفت وأوبن إيه آي.

أهلًا بكم في يوم سيتذكره العالم لفترة طويلة، أهلًا بكم في السادس من ديسمبر 2023، حيث

أطلقت جوجل النموذج الضخم الأول متعدد الوسائط: "جيميناي (Gemini).

قبل كل شيء، ما هو تعدُّد الوسائط؟

نحن نتحدث الآن عن شيء جديد من نوعه بمعنى الكلمة، نتحدث هنا عن النماذج اللغوية الكبيرة متعددة الوسائط. كي أبسّط عليك الأمر، دعني أعيد على مسامعك ماهية النماذج اللغوية الكبيرة التي أصبحت شيئًا تقليديًّا عفا عليه الزمن بعد هذا اليوم.

النماذج اللغوية الكبيرة والتي يرمز لها بـ LLM اختصارًا لـ Large Language Model، هي أحد تطبيقات الذكاء الاصطناعي المدرّبة على كمّ هائل من البيانات، والتي تكون عبارة عن نصوص ورموز.

أحدثت النماذج اللغوية الكبيرة (LLMs) مثل "شات جي بي تي 4" ثورة في طريقة تفاعلنا مع التكنولوجيا، تستطيع هذه النماذج أن تعالج اللغة الطبيعية وفهم وكتابة أنواع مختلفة من المحتوى الإبداعي بأسلوب أقرب إلى البشر منه إلى الآلة، كما يمكنها الترجمة بين الكثير من اللغات والإجابة عن الأسئلة، لكن كل هذا يسمح بفعله فقط بواسطة النصوص والرموز.

كان هذا أهم ما ميّز الأرباع الثلاث الأولى من عام 2023، أما الربع الأخير فقد شهد أخيرًا النماذج اللغوية الكبيرة متعددة الوسائط.

> "يستحق الكمبيوتر أن يطلق عليه "ذكي" إذا كان بإمكانه خداع إنسان في الاعتقاد بأنه إنسان آخر".. آلان تورينغ، أبو علم الحاسوب.

يعدّ النوع الأخير امتدادًا للنماذج اللغوية الكبيرة، لكنه يضيف إليها قدرات لمعالجة وفهم أشكال أخرى من البيانات، قد تتمثل في صور أو مقاطع فيديو أو ملفات صوتية.

ببساطة، يمكن اختصار الفارق بمثال بسيط، إذا سألت نموذج لغوي كبير كـ"جوجل بارد" قبل 6 ديسمبر/ كانون الأول الجاري عن أين يقع برج القاهرة؟ كان ليجيبك أنه في القاهرة، أما إذا سألته اليوم فسيجيبك بموقعه الجغرافي ونبذة صغيرة عنه وصورة أو مقطع فيديو، ليوضِّح لك كيف يبدو.

بطل قصتنا اليوم هو جيميناي (Gemini) -تكتب وتنطق هكذا-، وهو أول نموذج من النوع الحديث.

أهلًا بكم في عصر جيميناي

قبل أيام، أعلنت جوجل عن "جيميناي"، نموذج الذكاء الاصطناعي اللغوي الكبير الجديد، لكن بقدرات تجعله الأول من نوعه نموذجًا متعدد الوسائط، ما يعني أنه لا يمكنه فقط فهم ومعالجة وإنشاء النصوص، إنما يفعل الأمر ذاته مع الصور ومقاطع الفيديو والمقاطع الصوتية، ما يجعله كما تودّ جوجل أداة قوية وفعّالة للغاية، ومساعدًا غير محدود القدرات ليمكنه تقديم المساعدة الحقيقية للمستخدمين.

لماذا صنعت جوجل جيميناي؟

كان لجوجل 3 أهداف رئيسية أثناء عملها على "جيميناي"، أول هذه الأهداف هو التعددية، ليمكنه التعامل مع كافة أشكال المحتوى سواء المقروء أو المرئي أو المسموع، بالإضافة إلى الأكواد والكثير من أنواع البيانات الأخرى.

أما ثانيها فهو العمل بكفاءة على الهواتف الذكية، ليصبح سهل الوصول والاستخدام من قبل المستخدمين؛ وأخيرًا أن يكون مقاومًا لتغيّرات المستقبل، فهو قابل للتكيُّف والتعلُّم مع تقدم التكنولوجيا وزيادة حجم البيانات المدخلة، والذي أثبت بتطوره خلال فترة التدريب التي تولت مسؤوليتها شركة جوجل "ديب مايند"، القسم المختص بالذكاء الاصطناعي في جوجل.

حققت جوجل هذه الأهداف الثلاث في نموذجها الجديد، والذي سنتعرّف إليه في الأسطر التالية.

ما الذي يستطيع جيميناي فعله؟

"جيميناي" هو تحديث لمعمارية نموذج "بارد"، وذلك لتحسين قدراته وأدائه، وقد تمّ الإعلان عن 3 نسخ منه، وهي:

- **جيميناي آلترا**: النسخة الأكبر والأكثر تطورًا من النموذج اللغوي الكبير، والتي ستحتوي على الميزات كافة، والذي سيكون مناسبًا أكثر للاستخدام على مستوى الشركات، قد أعلنت جوجل أنه سيتوفّر للمستخدمين بداية من العام القادم.
- **جيميناي برو**: النسخة التي غالبًا ما ستصبح الأكثر استخدامًا، ذلك لأنها التي أصبحت متوفرة فور نهاية مؤتمر جوجل في 6 ديسمبر/ كانون الأول عبر نموذج جوجل الذكي للمحادثة "جوجل بارد"، وذلك في أكثر من 170 دولة على مستوى العالم، من بينها مصر والسعودية ولبنان والإمارات وغيرها، لكن فقط -حتى الآن- متوفِّر فقط باللغة الإنجليزية.
- **جيميناي نانو**: الإصدار الخفيف من نموذج "جيميناي"، وهو مخصص للاستخدام على الهواتف الذكية، من ناحية توفير الأداء الفعّال في الوقت الفعلي دون الحاجة للاتصال بالإنترنت، بالإضافة إلى إمكانية التواجد على الأجهزة محدودة الإمكانات، وهو حاليًّا يدعم مميزات مثل تلخيص التسجيلات الصوتية في تطبيقات المحادثة، مثل واتساب في الوقت الحالي على هواتف "بيكسل 8 برو"، بالإضافة إلى الردود الذكية عن طريق لوحة المفاتيح Gboard، وذلك بفضل شريحة Tensor G3، ومع الوقت نتوقع انتشار "جيميناي نانو" في بقية الهواتف العاملة بنظام آندرويد، بل قد يمتد ليعمل على الأنظمة غير الجوجلية أيضًا.

نحتاج لكي نفهم ما هو "جيميناي" أن نعرف أهم ما يمكنه فعله، وهو باختصار:

1. **الفهم الشامل**: يستطيع "جيميناي" معالجة وفهم المعلومات بأشكال مختلفة، ما يمكّنه من فهم العالم بطريقة أكثر شمولية.
2. **الاستدلال المعزز**: بفضل بنيته المتقدمة، يستطيع "جيميناي" التفكير وحلّ المشكلات بطريقة أكثر تطورًا، ما يؤدي إلى حلول أكثر ابتكارًا.
3. **الإبداع متعدد الأشكال**: تخيل آلة يمكنها إنتاج قصائد مستوحاة من لوحة فنية، أو تأليف سيمفونية بناءً على مقطع فيديو.
4. **لا مزيد من المعلومات المضللة**: تمكِّن معمارية "جيميناي" نموذج "بارد" من الوصول إلى المعلومات بشكل أفضل وتحليلها بدقة أعلى، ما يوفِّر لك تجربة دقيقة خالية من الأخطاء.
5. **المهنية**: من الكتابة الرسمية إلى التحليل العلمي، يقدم "جيميناي" تواصلًا على مستوى احترافي، ما يجعله أداة قيّمة للمحترفين في جميع المجالات.

جيميناي والنماذج الأخرى

فور الإعلان عنه، انطلق سيلٌ من المقارنات بين نموذج جوجل الجديد وأقوى نماذج شركة أوبن إيه آي المنافسة -شبه المملوكة لميكروسوفت-، "شات جي بي تي 4". مفاجأة، لن يخلو مقالنا هذا من مقارنة بين النموذجَين الأضخم في الوقت الحالي، لكن قبل ذلك دعني أوضِّح لك أمرًا آخر، وهو:

جوجل بارد قبل جيميناي

لعلّ أول مقارنة فكّرت بها كانت "جيميناي" و"جوجل بارد"، قبل مقارنته بـ"شات جي بي تي 4"، ذلك لأني كنت أنتظر إجابة عن "ما الذي قدمته جوجل لتحسِّن من نموذج "بارد" قبل أن تنافس "شات جي بي تي 4"؟"، وإليك الإجابة.

رغم قدراته الهائلة، كان يفتقر نموذج "بارد" قبل 6 ديسمبر/ كانون الأول عددًا من المزايا، كتركيزه على النصوص فقط ومعالجتها وإنشائها، وذلك حتى حصل على تحديث يمكنه من التعرُّف إلى الصور. لكن كان ذلك يأخذ وقتًا بعض الشيء، حيث كانت بنية بارد القديمة تعاني عند معالجة والاستجابة للطلبات المعقّدة.

كذلك كان "جوجل بارد" محدودًا، في النهاية لم يستطع تحقيق حلم آلان تورنغ بحاسوب لا تفرقه عن البشر، إضافة إلى تأليفه الكثير من الأحداث غير الحقيقية والخيالية، والاستدلال بها عندما يعجز عن البحث عن مثال إذا طلبت منه، وقد واجهت شخصيًّا هذا الأمر عدة مرات.

اختلف الأمر كثيرًا الآن، فمع إطلاق تحديث "جيميناي"، أصبحت قدرات "جوجل بارد"، والذي يمكن أن نطلق عليه الآن "واجهة المحادثة الخاصة بـ"جيميناي" أكثر قوة وفعالية عن ذي قبل، وذلك من خلال معالجة وإنشاء أنواع مختلفة من المحتوى، والترجمة بدقة وأكثر طلاقة، بالإضافة إلى قدرات كتابة الأكواد وتأليف المقطوعات الموسيقية وغيرها من الأشياء التفاعلية، بدقة وجودة عالية.

لهذا، يمكننا أن نعتبر أن جوجل قد نجحت بالفعل في تقديم شيء جديد في نموذجها السابق، وليس فقط محاولة المنافسة العمياء مع "شات جي بي تي 4"، من خلال اقتناص إحدى مميزاته والتعديل عليها، أما ما حدث في الناحية الأخرى فهو تفوق على كافة الأصعدة والنواحي.

جيميناي ضد شات جي بي تي 4

إنها المقارنة المرتقبة، بين أقوى أسلحة أوبن إيه آي، التي يرأسها منذ أيام فقط سام آلتمان.. أمزح، هو يرأسها منذ سنوات، لكني لا أجد مبررًا واقعيًّا لما حدث في الأسابيع السابقة سوى أنه مؤامرة للانقلاب على الإدارة السابقة.

هذا ليس موضوعنا، دعونا نعود إلى الحلبة، نعم، كنت أتحدث عن "جيميناي" ضد "شات جي بي تي 4"، ولكي أكون عادلًا سأستخدم في مثالي هذا النسخة الأعلى والأكثر تطورًا من "جيميناي": "جيميناي ألترا"، كونها الندّ الأفضل لـ"شات جي بي تي 4"، وإن كانت غير متوفّرة بعد للاستخدام، لكني سأستدل بما ذكرته جوجل في إعلانها عن "جيميناي".

نتائج اختبارات جيميناي ضد شات جي بي تي 4

> يتفوق جيميناي ألترا على جي بي تي 4.0 في معالجة النصوص – المصدر: جوجل ديب مايند

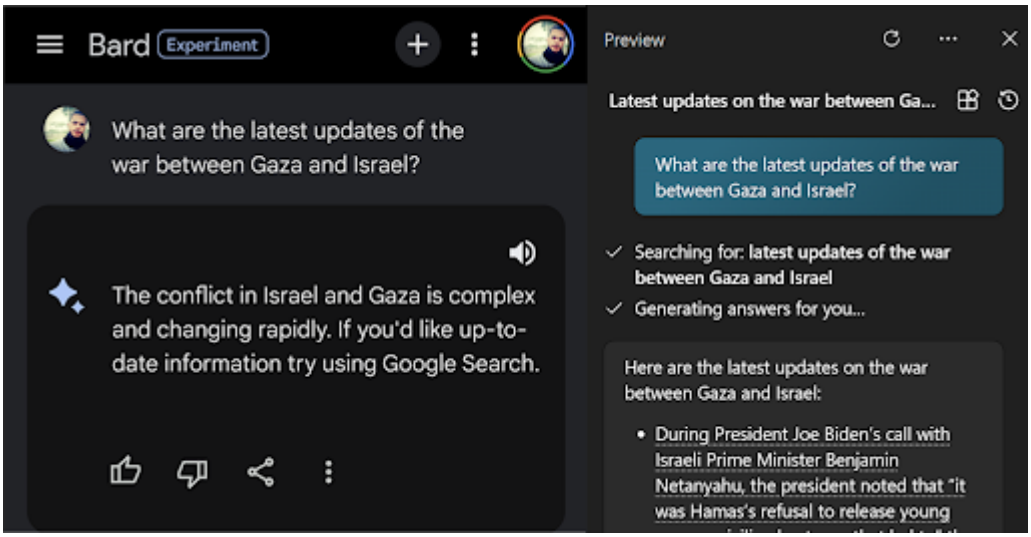
إذا دخلت إلى المدونة المخصصة لـ"جيميناي" على موقع "ديب مايند" الخاص بجوجل، ستجد استعراضًا للعديد من الجداول التي تحمل نتائج الاختبارات التي تمّت على كل من "شات جي بي تي 4" و"جيميناي ألترا".

أظهرت هذه الاختبارات تفوقًا واضحًا لنموذج شركة جوجل، كتفوقه بفارق 4.6% على "شات جي بي تي 4" في إتمام المهام المتعددة، والتفوق بفارق 7.7% في إنشاء أكواد لغة بايثون، وحصل على نسبة 94.4% في حل المسائل الحسابية، في حين حصل "شات جي بي تي 4" على 92% فقط، بالإضافة إلى التفوق بنسب متقاربة في العديد من النطاقات الأخرى كما تبيّن الصورة.

Capability	Benchmark Higher is better	Description	Gemini Ultra	GPT-4 API numbers calculated where reported numbers were missing
General	MMLU	Representation of questions in 57 subjects (incl. STEM, humanities, and others)	90.0% CoT@32*	86.4% 5-shot* (reported)
Reasoning	Big-Bench Hard	Diverse set of challenging tasks requiring multi-step reasoning	83.6% 3-shot	83.1% 3-shot (API)
	DROP	Reading comprehension (F1 Score)	82.4 Variable shots	80.9 3-shot (reported)
	HellaSwag	Commonsense reasoning for everyday tasks	87.8% 10-shot*	95.3% 10-shot* (reported)
Math	GSM8K	Basic arithmetic manipulations (incl. Grade School math problems)	94.4% maj@32	92.0% 5-shot CoT (reported)
	MATH	Challenging math problems (incl. algebra, geometry, pre-calculus, and others)	53.2% 4-shot	52.9% 4-shot (API)
Code	HumanEval	Python code generation	74.4% 0-shot (IT)*	67.0% 0-shot* (reported)
	Natural2Code	Python code generation. New held out dataset HumanEval-like, not leaked on the web	74.9% 0-shot	73.9% 0-shot (API)

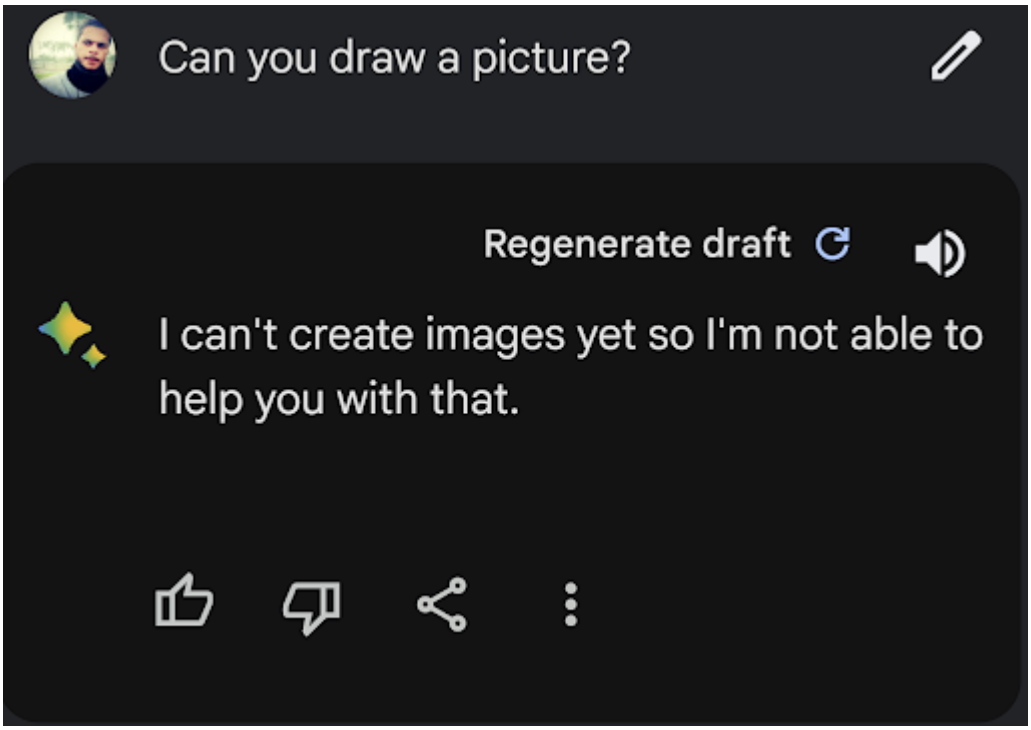
تجارب لكشف الاختلاف

حتى هذه اللحظة، لم يتمكن أحد من التحقق من هذه النسب التي أعلنتها جوجل دون السماح لأحد بمراجعتها، كون إطلاق النسخة ألترا من "جيميناي" مؤجّلة للعام المقبل، إلا أنه عند تجربة نسخة برو المتوفّرة بالفعل في نموذج المحادثة "بارد" ونسخة "شات جي بي تي 4" المتوفّرة من خلال أداة Copilot بسؤالهما عن آخر أخبار الحرب بين غزة و"إسرائيل"، ما زال يصرّ "بارد" على الإجابة بأن هذه الحرب معقدة، وينصح بالبحث في نتائج جوجل، في حين كانت إجابة "شات جي بي تي 4" عبارة عن نقاط فيها أحدث مستجدات الحرب، مع المصادر التي تمكّنك من الوصول إلى هذه المعلومات.



جوجل بارد مازال يعتبر الحرب أمرًا شائكًا لا يجوز الحديث فيه

وعند سؤال "بارد" عن إمكانية رسم الصور، ما زال يردّ بعدم إمكانيته، في مقابل قدرة "شات جي بي تي 4" على تنفيذ المهمة ذاتها دون عناء، ولكن هنا يوجد اختلاف هام يوضِّح لك أكثر الفارق بين ال LLM وال MMLLM.



"جيميناي برو" المدمج في "بارد" يفتقر لقدرات إنشاء وتصميم الصور

النموذج الأقرب للذكاء العام

في الحالة الأولى، والتي يمثِّلها هنا نموذج "شات جي بي تي 4" والنسخة قبل "جيميناي" من "بارد"، كان يتفوق الأول في قدرته على رسم الصور باحترافية كبيرة، لكنه كان يعتمد في هذا على إرسال الطلبات إلى نموذج DALL-E 3، النسخة الفنية من نماذج أوبن إيه آي الذكية، والذي كان يمكنه الرسم بكل سهولة.

أما على الجانب الآخر، فإن ما سنراه عمّا قريب من خلال "جيميناي ألترا" هو نموذج قادر على استقبال الأوامر الصوتية أو النصية أو المرئية، وتنفيذها بما يتضمن رسم وتصميم الرسومات دون الحاجة لأي دعم خارجي، فهو أقرب إلى الذكاء الاصطناعي العام (AGI) منه إلى التقليدي، وإذا لم تسمع عن هذا المصطلح من قبل، قد أتحدث عنه في مقالة لاحقة، لكن يلخصه البعض بأنه نهاية البشرية!

بخلاف هذا وذاك، فإن صديقنا الأقدم "شات جي بي تي 4" مازال متفوقًا -ولو بشكل مؤقت- في بعض النقاط، كمجتمع الدعم والأدوات الموجود لخدمة مستخدميه، لكن ذلك يعود إلى تواجده على الساحة وتوفّر ال APIs الخاصة به منذ فترة طويلة، وهو شيء لا يصعب على جوجل عمله في

غضون أسابيع من بعد إطلاقها لـ"جيميناي ألترا".

ومن ناحية أخلاقية أكثر من كونها تقنية، فإن "شات جي بي تي 4" يمتاز بالشفافية عن نموذج جوجل، حيث إن الأخير وصانعيه رفضوا -حتى لحظة كتابة هذا المقال- الردّ على أي أسئلة تخص كيفية جمع البيانات التي تمّ تدريب "جيميناي" عليها.

مضيئًا للأنظمة، موحشًا للبشر

إذا كنت قد استخدمت نموذج "بارد" مؤخرًا، ستجد رسالة جديدة باللون الأزرق، مكتوب عليها أن المحادثات التي تتم بينك وبين "بارد" أصبحت تراجَع من قبل البشر، مع تنويه بعدم مشاركة أية بيانات لا تودّ أن تراجَع من قبل فريق جوجل.

هكذا هو الأمر، فمع التقدّم تزداد المسؤولية، ويتطلب الأمر تدخلًا بشريًّا كي لا تخرج الأمور عن السيطرة.

يعدّ "جيميناي" علامة فارقة في قدرات النماذج اللغوية الكبيرة، والتي انتقلت من أن تصبح مجرد نماذج لغوية إلى نماذج متعددة الوسائط، يفتح هذا التحديث آفاقًا جديدة لاستخدام الذكاء الاصطناعي في مجموعة متنوعة من المجالات، بما في ذلك التعليم والصحة والرعاية والخدمات المالية.

مع ذلك، فإن تحديث "جيميناي" يعيد بعض التحديات الهامة. فما هي الأخلاقيات المتعلقة باستخدام هذه النماذج؟ كيف يمكننا التأكد من استخدامها بطريقة مسؤولة وأخلاقية؟ هل يمكن للذكاء الاصطناعي أن يكون أخلاقيًّا؟ خاصة مع تفوق هذا النموذج الأخير لأول مرة على الخبراء البشريين في قدرات فهم اللغة المتعددة الضخمة (MMLU)، وهي إحدى أكثر الطرق شيوعًا لاختبار معرفة وقدرات حل المشكلات لنماذج الذكاء الاصطناعي، والتي تفوّق فيها الإنسان سابقًا على "شات جي بي تي 4" بفارق 3.4%.

هناك العديد من وجهات النظر المختلفة حول هذه القضية، يعتقد بعض الخبراء أن الذكاء الاصطناعي لا يمكن أن يكون أخلاقيًّا حقًّا، لأنه مجرد أداة يمكن استخدامها للأغراض الخيرية أو الشريرة، ويعتقد آخرون أن الذكاء الاصطناعي يمكن أن يكون أخلاقيًّا، لكن فقط إذا تم تصميمه واستخدامه بطريقة أخلاقية.

يطرح تحديث "جيميناي" أيضًا أسئلة حول طبيعة الذكاء البشري، إذا أضحت هذه النماذج قادرة على فهم وإنتاج اللغة البشرية، فهل هذا يعني أنها ذكية مثل البشر؟ أو في مرحلة ما قد تكتسب الوعي؟ أم أنها ما زالت لا تتمتع بالقدرة نفسها على التفكير النقدي والإبداعي؟

لا توجد إجابات سهلة لهذه الأسئلة، ومع ذلك إن تحديث "جيميناي" يسلط الضوء على أهمية

التفكير في هذه القضايا بشكل جدّي، خاصة مع وجود هذا الكمّ الهائل من القدرات المملوكة لشركة واحدة تدعى جوجل، الشركة التي -وبكل تأكيد- تدري ما تفعله جيدًا، وما يدرّ عليها الربح بشكل مستمر، كما حدث عندما ارتفعت أسهمها بنسبة 5% بعد إطلاق "جيميناي" بساعات قليلة.

رابط المقال : https://www.noonpost.com/184862/