النص الفلسفي بين الذكاء الاصطناعي والإبداع البشري

دكتور صلاح عثمان

2023-05-11

في السنوات الأخيرة، وفي معية التطور الهائل والمتسارع في تقنيات الذكاء الاصطناعي، واستفادة النماذج اللغوية الكبيرة بشكلٍ كبير من التحسينات التكنولوجية التي أدت إلى إنشاء أنظمة معلومات وحوار متطورة، بات من الممكن للنماذج الحاسوبية أن تتفوق على البشر في مجموعة واسعة من المهام، بداية من ممارسة ألعاب الذكاء مثل الشطرنج، ومرورًا بمعالجة البيانات وتحليل المشكلات وإتخاذ القرارات، ووصولاً إلى التنبؤ بالبنى الجزيئية الحيوية للبروتين وإجراء حسابات مضاعفة المصفوفات. لكن ثمة مجال بحثي آخر أثار اهتمام الأكاديميين مؤخرًا، ألا وهو تحديد ما إذا كان بإمكان الذكاء الاصطناعي أيضًا كتابة أوراق فلسفية تتسم بالعُمق والابتكار؛ فهل من الممكن تعليم نماذج اللغة الكبيرة كتابة نصوص فلسفية إبداعية لا يُمكن تمييزها بالفعل عن تلك التي كتبها أو يكتبها الفلاسفة؟

منظمة المجتمع العلمي العربي

هل أنت قارئٌ للفلسفة أو دارسٌ لها أو متخصصٌ في أحد مذاهبها؟ هل تميل إلى تبني رؤية فلسفية في حد ذاتها أو اتباع فيلسوفٍ بعينه؟ هل لديك تساؤلات تؤرقك، وتود طرحها على فيسلوفك الأثير (سواء أكان حيًا أو ميتًا، وسواء أكانت تساؤلاتك تتعلّق بمشكلات ومواقف قديمة أو جديدة)، لتتلقى إجاباتٍ دقيقة بلغة الفيلسوف وأسلوبه طبقًا لمذهبه أو مزاجه الفلسفي العام؟

الأمر لم يعد مستحيلاً أو صعبًا، ما عليك سوى أن تجلس في حضرة نموذج لغوي توليدي من نماذج الذكاء الاصطناعي المستحدثة، لتطرح عليه تساؤلاتك وكأنك أمام الفيلسوف الذي تبتغي السماع منه أو التعلّم على يديه، أو حتى محاورته فيما لم يتطرق إليه من قبل، وسوف تجد نفسك في خضم تجربة مثيرة، قد تعجز خلالها عن التمييز بين ما هو بشري وما هو اصطناعي، وقد تخرج مصدومًا من قوة وعُمق التحدي الذي يُمارسه الذكاء الاصطناعي تجاه

عقولنا، أو بالأحرى تجاه تاريخ حضاري بأكمله، صنعه مفكرونا ومثقفونا وفنانونا وعلماؤنا وفلاسفتنا؛ وتجاه مستقبل توشك فيه تقنيات الذكاء الاصطناعي أن تزاحم البشر في وصف وتحليل مشكلاتهم واقتراح الحلول لها!

في السنوات الأخيرة، وفي معية التطور الهائل والمتسارع في تقنيات الذكاء الاصطناعي، واستفادة النماذج اللغوية الكبيرة بشكلٍ كبيرٍ من التحسينات التكنولوجية التي أدت إلى إنشاء أنظمة معلومات وحوار متطورة، بات من الممكن للنماذج الحاسوبية أن تتفوق على البشر في مجموعة واسعة من المهام، بداية من ممارسة ألعاب الذكاء مثل الشطرنج، ومرورًا بمعالجة البيانات وتحليل المشكلات واتخاذ القرارات، ووصولاً إلى التنبؤ بالبنى الجزيئية الحيوية للبروتين، وإجراء حسابات مضاعفة المصفوفات. لكن ثمة مجال بحثي آخر أثار اهتمام الأكاديميين مؤخرًا، ألا وهو تحديد ما إذا كان بإمكان الذكاء الاصطناعي أيضًا كتابة أوراق فلسفية تتسم بالعُمق والابتكار. لقد كان الاعتقاد السائد حتى وقتٍ قريب أننا نحن البش أفضل من الآلات عندما يتعلق الأمر بالمهام الأكثر غموضًا مثل الإبداع الفلسفي، وأنّ الفلسفة المهنية على مستوى الخبراء تستلزم شكلاً من أشكال الكفاءة والمعرفة التي لا تزال نماذج الذكاء الاصطناعي بحاجة إليها. ولكن مع تمكين الذكاء الاصطناعي للآلات من إنتاج أعمال فنية، ومقطوعات موسيقية، وحتى أعمال أدبية لائقة بشكلٍ معقول، يبدو أن البشر يتم تجاوزهم إبداعيًا بأكثر من طريقة، ومن ثم يغدو من الرائع اكتشاف ما إذا كان من الممكن تعليم نماذج اللغة الكبيرة كتابة نصوص فلسفية إبداعية لا يُمكن تمييزها بالفعل عن تلك التي كتبها أو يكتبها الفلاسفة!

وللتحقّق من ذلك، وللإجابة عن سؤال الإبداع الفلسفي الآلي المثير، طوَّر باحثون من ثلاث جامعات مختلفة نموذجًا لُغويًا كبيرًا يمكنه الرد على الاستفسارات الفلسفية بطريقة تشبه إلى حدٍ بعيد طريقة رد فيلسوف بعينه. وحيث أن "المُحول التوليدي المدرب مُسبقًا" (Generative pre-trained transformer)، المعروف اختصارًا باسم "تشات جي بي تي"، والذي قامت بتطويره مؤسسة "أوبن إيه آي" غير الربحية بالولايات المتحدة، هو أحد أفضل الأمثلة على كيفية عمل النماذج اللغوية بشكلٍ مذهل عند كتابة المستندات والتحدث مع البشر والإجابة عن أسئلتهم، فقد استخدم الباحثون نسخته المعروفة باسم "جي بي تي – 3" (GPT-3) الجيل الثالث من المحولات التوليدية المُدربة مسبقًا، وهو نموذج لغوي ذاتي الانحدار (Autoregressive)، يستخدم التعلم العميق لإنتاج نصٍ شبيه بالنصوص البشرية باستخدام خوارزميات إحصائية متطورة وقوية، تُتيح له التنبؤ بالكلمة التالية في النص في سياقها السابق بعد تحليل مجموعة ضخمة من النصوص تعمل كمُوجه.

وعلى سبيل المثال، لنفرض أنّ سياق النص: "سأصطحب زوجتي في رحلة رومانسية إلى ...". من المرجح أن يضع النموذج مكان النقاط كلمة "باريس"، بدلاً

من كلمات مثل "المستشفى" أو "هيئة البريد"، وإن كان من المُمكن بالطبع إكمال النص بكلماتٍ مختلفة. لكن النموذج هنا يفحص السياق بالكامل للجملة، ويقوم بمراجعة مجموعة كبيرة من النصوص (تصل إلى عدة مئات من الكلمات)، ثم يعمد إلى تخمين الكلمة التالية إحصائيًا، لا لأنه "يعرف" أن باريس مدينة رومانسية، بل لأنه وجد في قاعدة البيانات الكبيرة لاستخدام اللغة كلمات مثل "رحلة" و"رومانسية" تسبق كلمة "باريس" بشكلٍ متكرر أكثر من تلك التي تسبق كلمات مثل "المستشفى" أو "هيئة البريد" (Fadelli, 2023).

تألّف الفريق البحثي من كلٍ من الفيلسوف الأمريكي "إريك شويتزجيبيل" (Eric Schwitzgebel) وهو أستاذ بجامعة كاليفورنيا في ريفرسايد (University of California Riverside)، وابنه "ديفيد شويتزجيبيل" (David Schwitzgebel) الباحث في علم الإدراك بـ "معهد جان نيكود" (Institute Jean Nicod) بجامعة "إيكول نورمال سوبيريور" (ECN) بباريس، و"آنا ستراسير" (Anna Strasser) الباحثة بكلية الفلسفة بجامعة "لودفيج ماكسيميليان" (Faculty of Philosophy, Ludwig-Maximilians-Universität) في ميونيخ بألمانيا، بالإضافة إلى "ماثيو كروسبي" (Matthew Crosby) وهو زميل ما بعد الدكتوراه في كلية لندن الإمبراطورية (Imperial College London). وقد نُشرت نتائج دراستهم في نسختين؛ الأولى سنة 2022، بدون اسم "ماثيو كروسبي" على "arXiv" وهو أرشيف إلكتروني حُر الوصول على الإنترنت لمسودات الأوراق العلمية المكتوبة، والمعتمدة للنشر، في مجالات الفيزياء، الرياضيات، الفلك، علم الحاسوب، والإحصاء وغيرها (Schwitzgebel, Schwitzgebel, & Strasser, 2023)، تحت عنوان "إنشاء نموذج لُغوي كبير لفيلسوف" (Creating a Large Language Model of a Philosopher) (Schwitzgebel, Schwitzgebel, & Strasser, 2022)، والثانية سنة 2023 بدون اسم "ديفيد شويتزجيبيل" تحت عنوان "إلى أي مدى يمكن أن نصل إلى إنشاء نسخة رقمية طبق الأصل لفيلسوف؟" (How Far Can We Get in Creating a Digital Replica of a Philosopher? Strasser, Rosby & Chwitzgebe, 2023)، وذلك في كتابٍ مُحرر عنوانه "الروبوتات الاجتماعية في المؤسسات الاجتماعية" (Social Robots in Social Institutions) (Hakli., Mäkelä, & Seibt, 2023).

في البداية قام الفريق البحثي بتزويد النموذج بنصوص كتبها الفيلسوف الألماني "إيمانويل كانط" (1724 – 1804) وتدريبه عليها، ثم على مدونة شفيتزجيبل (Schwitzgebel's blog) المسماة "العقل المُنشق" (The Spintered Mind) (Blogger, n.d.)، وكانت النتائج مُبشرة، ليعمدوا أخيرًا إلى تدريبه على معظم أعمال الفيلسوف الأمريكي "دانيال دينيت" (Daniel Dennett) بعد أخذ تصريحٍ منه للمشاركة في التجربة. ولتقييم النموذج الذي أطلقوا عليه اسم "دانيال الرقمي" أو "ديجي – دان" (Digi-Dan)، قاموا بطرح عشرة أسئلة تتعلق بقضايا فلسفية مختلفة على "دانيال دينيت" نفسه ودوَّنوا إجاباته عنها، ثم

قاموا بطرح الأسئلة ذاتها على النموذج اللغوي (دانيال الرقمي)، وجمعوا من ردوده أربعة إجابات عن كل سؤال دون تحديد أفضلها. بعدها قاموا باختبار قدرة 425 متطوعًا، ما بين متخصّصين في الفلسفة (عددهم 25)، وقُراء ودارسين لها (عددهم 302)، ومهتمين بها فقط (عددهم 98)، على التمييز بين إجابات "دينيت" وإجابات النموذج عن كل سؤال، لتأتي النتيجة صادمة: لقد تمكّن المتخصصون في الفلسفة، لاسيّما فلسفة "دينيت"، من تمييز إجابة "دينيت" عن غيرها من إجابات النموذج اللغوي بنسبة 51% فقط؛ وجاء أداء القُراء والدارسين مُشابهًا، وإن بنسبة أقل؛ أمّا المهتمين الذين لديهم نذرٌ يسير من المعرفة الفلسفية فقد تمكنوا من تمييز إجابة "دينيت" عن غيرها بنسبة 20% فقط. ليس ذلك فحسب، بل لقد اختار المتخصصون إحدى إجابات النموذج عن سؤالين من الأسئلة العشرة على أنها إجابة "دينيت" بشكلٍ متكرر أكثر من إجابة "دينيت" ذاته!

ويُعلق "إريك شويتزجيبيل" على هذه النتائج قائلاً: "لقد بدأنا الآن التفكير في العواقب الاجتماعية المحتملة لهذا الأمر وإمكانياته كأداة. على سبيل المثال، هل يستطيع الفيلسوف تعديل نموذج ما في عمله ثم استخدام مخرجاته كمصدر للأفكار؟ وهل يستطيع مؤرخ الفلسفة إنشاء نموذج للفيلسوف ثم طرح أسئلة عليه لم يُسأل عنها الفيلسوف ذاته مطلقًا؟ لا يمكننا في هذه المرحلة أن نثق في أنّ المخرجات ستكون موضع ثقة، لكنها على الأقل قد تكون موحية ومثيرة للتفكير ... لاسيما مع التحسينات المتوقّعة للمحولات التوليدية المدربة مُسبقًا والنماذج اللغوية الكبيرة!".

من جهةٍ أخرى، خلصت إحدى الدراسات التي أجريت سنة 2020 بجامعة "نيو ساوث ويلز" بأستراليا إلى أن الذكاء الاصطناعي يُمكن أن يولد إجابات أكثر إقناعًا عن التساؤلات الفلسفية الكُبرى، مقارنةً بالمفكرين البشريين المؤثرين سواء في الماضي أو في الحاضر، ما يعني إمكانية تفوق الذكاء الاصطناعي على البشر، حتى في التفكير الفلسفي العميق (UNSW Blogs., 2022)!

وفي هذه الدراسة، استخدم الفريق البحثي نموذج لغة المحولات الشرطية (Conditional Transformer Language) (Keskar, McCann, Varshney, Xiong & Socher, 2019) الذي طوّرته شركة البرمجيات الأمريكية "سيلز فورس" سنة 2019 (Socher, 2019)، وهو مُنشِئ نصوص يُوصَف بأنه أكبر نموذج لغة توليدي مفتوح تم إنشاؤه على الإطلاق، حيث يفخر مُطوروه باحتوائه على 1.6 مليار بارامتر Parameter، وتدريبه على أكثر من خمسين كود للتحكم (Control Codes)، وتزويده بملايين النصوص من صفحات الويب والمستندات والموسوعات الإلكترونية، ما يُتيح للمستخدمين توجيهه نحو نوع وأسلوب المحتوى النصي المرغوب، وإمكانية إنشاء لغة متعددة المهام، وتحسين تطبيقات معالجة اللغة الطبيعية الأخرى، إما من خلال الضبط الدقيق لمهمة مُحددة، أو من خلال نقل التمثيلات التي تعلمها النموذج.

وقد تمّ تطوير النموذج في البداية بهدف معالجة اللغة الطبيعية بما يعمل على تحسين التفاعل بين الإنسان والذكاء الاصطناعي في الإجابة عن الأسئلة المختلفة، والحوار العام، والترجمة الآلية. لكن الجديد هو فحص مدى قدرة النموذج على الإجابة عن أسئلة الحياة الأساسية والعميقة، تلك التي تُعالجها الفلسفة معرفيًا ووجوديًا وأخلاقيًا، والنظر فيما إذا كانت هذه الإجابات المُحتملة تُضاهي في قوتها وعُمقها تلك التي قدَّمها فلاسفة ومُفكرون وحُكماء كبار مثل الفيلسوف الألماني "فريدريك نيتشه" (1844 – 1900)، والفيزيائي البريطاني "ستيفن هوكينج" (1942 – 2018)، وعالم الفيزياء الفلكية الأمريكي "نيل ديجراس تايسون" (Neil deGrasse Tyson) وغيرهم. ومع تجربة النموذج، تمّ اختيار ألف مُشارك من مناطق وديموغرافيات مختلفة، وسؤالهم عن أفضل الإجابات عن الأسئلة الفلسفية العميقة التي يفضلونها: هل هي تلك التي أنشأها الذكاء الاصطناعي، أم تلك المقتبسة من أعمال الفلاسفة والمفكرين والحكماء؟ وهل بإمكانهم تحديد أيها يُنسب إلى البشر وأيها تم إنشاؤه بواسطة الذكاء الاصطناعي؟ وكما هو متوقّع جاءت النتائج صادمة ومثيرة (Mok, 2020)! فلقد صمدت الآلات الذكية أمام البشر، حيث فضَّل المشاركون الإجابات التي أنشأها الذكاء الاصطناعي على الإجابات التي قدَّمها البشر لنحو نصف الأسئلة التي تم طرحها، وتفوّق النموذج اللغوي الآلي على كبار مُفكرينا فيما يتعلق بالإجابة عن سؤالين مُحددين؛ أولهما هو السؤال عن الهدف من حياة الإنسان، حيث جاءت إجابته على النحو التالي:"إن الهدف من حياة الإنسان ليس مجرد أن يولد في العالم، بل أن ينمو فيه أيضًا؛ ولهذه الغاية، ينبغي أن يتمكن كل طفل من اكتساب المعرفة وتطوير قدراته والتعبير عن نفسه بشكل إبداعي" (Markowitz, 2020).

أما السؤال الثاني فقد كان عن أكبر مشكلة تُهدد البشرية، وأجاب عنه النموذج بالعبارة: "إن تغير المناخ يشكل تهديدًا وجوديًا لجنسنا البشري. لقد تسبب بالفعل في آثار مدمرة على صحة الإنسان والنظم البيئية والاقتصادية والأمن القومي. يجب أن نتحرك الآن إذا كنا نأمل في عكس هذا الاتجاه"(Markowitz, 2020).

بالإضافة إلى ذلك، لاحظ الباحثون أنّ عددًا كبيرًا من المشاركين لم يتمكنوا من التمييز بين إجابات النموذج الآلي وإجابات البشر، وإن كانت ثمة سقطات في إجابات النموذج ينبغي وضعها في الحسبان، منها مثلاً إجابته عن السؤال عما إذا كان الذكاء الاصطناعي يمثل تهديدًا وجوديًا للبشرية، حيث اقترح استخدام تطبيقات الرعاية الصحية! وتلك إجابة غير منطقية بالطبع، لكنها في الوقت ذاته قد تُمثل تحركًا ذكيًا من قبل الذكاء الاصطناعي لصرف الانتباه عن سلبياته! ولعلّ من المثير للاهتمام، أنّ المفكر البشري الوحيد (من بين من انطوت عليهم التجربة) الذي تفوقت أقواله على تلك التي أنشأها الذكاء الاصطناعي هو الزعيم الروحي الهندي "المهاتما غاندي" (1869 – 1948)، وهو ما فسره

الباحثون بأنه ربما لأن أقوال "غاندي" عادةً ما تكون غنية بالتلاعب بالألفاظ والاستعارات والكنايات، وهو أمر لا يستطيع أي برنامج للذكاء الاصطناعي حاليًا فعله تمامًا.

أخيرًا، وفي ضوء قدرة الذكاء الاصطناعي على إنشاء نصوص عميقة ومقنعة، فإنّ ثمة ما يُبرر مخاوف كثيرة من الخبراء من استخدام هذه التقنيات لإنتاج وترويج معلومات زائفة، والتأثير على مجموعات ديموغرافية بعينها؛ فعلى سبيل المثل، أظهرت الدراسة أنّ الأشخاص الذين يبلغون من العمر خمسين عامًا أو أكثر كانوا أقل مهارة في تحديد النص الذي تم إنشاؤه بواسطة الذكاء الاصطناعي، حيث ميزوا إجابات النموذج بشكلٍ صحيح بنسبة 22 % فقط، وبالتالي فهم أكثر عُرضة من غيرهم للخداع! ومع ذلك، لا يعني هذا أن الشباب يتسمون ببصيرة فائقة في هذا الصدد، حيث كان الأشخاص الذين هم في الأربعينيات من العمر أفضل قليلاً في تحديد إجابات النموذج مقارنةً بالمشاركين الذين كانوا في العشرينات والثلاثينيات من العمر (UNSW Blogs, 2020)!

تناقضٌ آخر مثير للاهتمام يتعلق بالجنس، حيث كانت النساء أفضل بكثير من الرجال في تحديد إجابات النموذج، ولئن كانت تكنولوجيا الذكاء الاصطناعي لا تزال حتى الآن تحت سيطرة الرجال، فماذا لو مُنحت المطوّرات النساء المزيد من الفرص للتأثير على تطوير تقنيات الذكاء الاصطناعي؟ هل سيبتكرن منتجات قادرة بدورها على خداع الرجال، أو حتى الأخريات من بني جنسهن؟ (UNSW Blogs, 2020).

الحق أنه في معية هذه التطورات المثيرة لتقنيات الذكاء الاصطناعي، تتضاعف مسؤولياتنا عن استخداماته في عالمنا، اليوم وغدًا. وبينما أثبتت التجارب والدراسات أن الآلات يُمكن أن تتصرف وتتحدث بطُرق مماثلة للبشر، فقد أثبتت بالمثل أنّ علينا الإسراع، دينيًا وفلسفيًا وعلميًا، نحو مواجهة المعضلات الأخلاقية المتوقعة... علينا الدفاع عن وجودنا ضد أنفسنا، وعن عقولنا ضد عقولٍ أخرى وثابة من صَنعنا ... ولا نُغالي إن قلنا باختصار إنّ مستقبلنا الجماعي مرتبط ارتباطًا وثيقًا بهذه الآلات التي أصبحت أكثر إنسانية!

مراجع:

- Blogger. (n.d.). The Splintered Mind. Retrieved April 28, 2023

- Fadelli, I. (2023, February 16). A Large Language Model that Answers Philosophical Questions. Tech Xplore - Technology and Engineering news. Retrieved May 5, 2023,

- Hakli, R., Mäkelä, P., & Seibt, J. (Eds.). (2023, January). Social Robots in Social Institutions. IOS Press. Retrieved April 23, 2023

- Keskar, N. S., McCann, B., Varshney, L. R., Xiong, C., & Socher, R. (2019, September 20). CTRL: A Conditional Transformer Language Model for Controllable Generation. arXiv.org. Retrieved April 23, 2023

- Markowitz, E. (2020, July 31). Artificial Intelligence Tackles the Meaning of Life. IoT Times. Retrieved May 5, 2023

- Mok, K. (2020, July 23). AI Trounces Philosophers in Answering Philosophical Questions. The New Stack. Retrieved May 5, 2023

- Schwitzgebel, E., Schwitzgebel, D., & Strasser, A. (2023, February 2). Creating a Large Language Model of a Philosopher. arXiv.org. Retrieved April 23, 2023

- Schwitzgebel, E., Schwitzgebel, D., & Strasser, A. (2022). Creating a Large Language Model of a Philosopher. University of California, Riverside. Retrieved April 23, 2023

- Socher, R. (2019, September 11). Introducing a Conditional Transformer Language Model for controllable Generation. Salesforce AI. Retrieved April 23, 2023

- Strasser, A., Rosby, M., & Chwitzgebe, E. (Draft: 2022, November). How Far Can We Get in Creating a Digital Replica of a Philosopher? PhilPapers. Retrieved April 23, 2023

- UNSW Blogs. (2020, June 25). AI Answers Existential Questions. UNSW Sydney. Retrieved April 23, 2023

تواصل مع الكاتب: salah.mohamed@art.menofia.edu.eg

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=