

العصر الجديد للذكاء الاصطناعي العربي

جمال مراد قيس

2025-09-29

في السنوات الأخيرة، أخذت اللغة العربية تحظى باهتمام متزايد في بحوث الذكاء الاصطناعي، لكن ما كان يُفتقد هو وجود نماذج لغوية كبيرة "Large Language Models – LLMs" متخصصة بالعربية ومبنية بمنهجية بحثية رصينة تستجيب لخصوصية اللغة من حيث البنى المعجمية والنحوية واللهجات.

ومع تصاعد الجهود البحثية في هذا المجال، ظهرت مؤخراً ما يمكن تسميته "أرضية عربية للنماذج اللغوية الكبرى Arabic LLMs أو ALLMs، وهي مجموعة متنامية من المشاريع والدراسات التي تحاول سد الفجوة بين النماذج اللغوية العامة والنماذج التي تفهم العربية بعمق.

يُعد مشروع "ALLaM" واحداً من أبرز هذه المشاريع العربية الحديثة؛ حيث طور فريق من الباحثين نماذج لغوية كبيرة تدعم اللغة العربية والإنجليزية معاً، مع مراعاة التوفيق اللغوي ونقل المعرفة بين اللغتين. استخدم الباحثون بيانات مزيجة من العربية والإنجليزية في تدريب النماذج، بالإضافة إلى استخدام بيانات مترجمة ومحاكاة بشرية لتحسين أداء النموذج بالعربية دون التضحية بالكفاءة في الإنجليزية. وقد حققت هذه النماذج أداءً متقدماً في مقاييس عربية مثل MMLU Arabic وACVA وغيرها.

إلى جانب ALLaM، ظهرت مشاريع أخرى مثل ArabianGPT التي تتبع نهجاً أكثر توجيهاً نحو العربية، حيث ركزت على بناء معمارية متخصصة للنصوص العربية مع استخدام وحدة ترميز "tokenizer" تلائم البنية الصرفية واللفظية للغة. في اختبارات التطبيق، أظهرت النسخ الدقيقة من هذا النموذج تحسناً في مهام مثل تحليل المشاعر والتلخيص مقارنة بالنماذج العامة التي لا تأخذ خصوصية العربية بعين الاعتبار.

كما ظهر النموذج AIN Arabic INclusive multimodal model مؤخراً ليجمع بين النص والصورة في إطار ثنائي اللغة عربية وإنجليزية. يهدف هذا النموذج إلى معالجة محتوى بصري عربي وفهمه ضمن السياق اللغوي العربي، مما يفتح

آفاقاً جديدة لتطبيقات مثل تحليل الصور للغة العربية أو ترجمة النصوص المصوّرة.

لكن التحديات ما تزال ضخمة أمام الباحثين العرب في هذا المجال. أولها قلة البيانات العربية عالية الجودة مقارنة بالإنجليزية، مما يجعل عملية التدريب المعقدة صعبة. دراسات مسحية في المشاريع المتخصصة تقول إن تنوع اللهجات، الكلمات الغامضة، والتداخل بين العامية والفصحى كلها عوامل تشكّل عبئاً إضافياً على بناء نماذج عربية متينة.

كما أن التقييمات الواقعية للنماذج العربية غالباً ما تفتقر إلى معايير موحدة، خصوصاً فيما يتعلق بالمقاييس الثقافية والدقة المفهومية. في هذا السياق، تم اقتراح معيار "AlGhafa" كمقياس اختباري للنماذج العربية لتقييم قدرتها على فهم الأسئلة متعددة الاختيارات بالاعتماد على محتوى عربي متنوّع، واختبار جودتها في السياق المحلي.

في ضوء هذه التطورات، يبدو أن أرضية البحث العربية في نماذج اللغة الكبيرة تتحول من حالة التأسيس إلى حالة التنافس والابتكار. ما كان يُنظر إليه قبل سنوات كمشروع طموح أصبح اليوم حقيقة تتبلور عبر نماذج تعمل، بيانات عربية متجددة، وبنى بحثية محلية. ولكن النجاح الكامل يتطلب مزيداً من التعاون بين الجامعات، المراكز البحثية، الصناعة، وتمويل مستدام لدعم البنى التحتية الحوسبية الضخمة.

أبرز ملامح أهمية هذه الأرضية البحثية

1- تعزيز الهوية اللغوية: تمكين اللغة العربية من أن تكون لغة إنتاج معرفي رقمي، وليس مجرد لغة ترجمة أو تدريب ثانوي. 2- توسيع نطاق البحث العلمي: إتاحة موارد ونماذج متخصصة بالعربية للباحثين والمطورين، ما يفتح المجال لمشاريع أكاديمية وتطبيقات عملية جديدة. 3- دعم الابتكار المحلي: تشجيع إنشاء شركات ناشئة عربية تعتمد على تقنيات الذكاء الاصطناعي اللغوي المحلي بدلاً من النماذج الأجنبية فقط. 4- سد الفجوة التقنية: المساهمة في تقليل الفجوة بين التقنيات العالمية المتقدمة والاحتياجات المحلية للمجتمعات العربية. 5- تمكين التطبيقات المتخصصة: تطوير أدوات معالجة اللغة العربية في مجالات التعليم، القانون، الإعلام، والخدمات الحكومية، بما يخدم المتحدثين باللغة العربية مباشرة.

في النهاية، يبقى الهدف الأكبر هو أن تصبح اللغة العربية ليست مجرد موضوع يُدرّس للنماذج الإنجليزية، بل أن تكون اللغة الأساسية التي تولد منها النماذج، بحيث تُفهم وتُنشج بها الأفكار والابتكارات بدقة وسلاسة، وتخدم أكثر من 400 مليون متحدث عربي بطريقة جديدة بالزمن الرقمي.

المصادر

-1 "Large Language Models for Arabic and English", M Saiful Bari et al
-2 [A Survey of Large Language Models for Arabic Language and its](#)
[ArabianGPT: Native Arabic GPT-based Large Language](#)
-3 ["Dialects](#)
[AlGhafa](#)
-5 ["The Arabic INclusive Large Multimodal Model](#)
-4 ["Model](#)
["Evaluation Benchmark for Arabic Language Models](#)

تواصل مع الكاتب: mohamedmouradgamal@gmail.com

[/https://arsco.org/articles/article-detail-47306](https://arsco.org/articles/article-detail-47306)