

# نماذج اللغة الكبيرة: دمج الخبرة يحسن التشخيص النفسي بدقة أعلى

تأليف

مدرس الدكتور محمد لوتي

فبراير 10, 2026

اقتبس من هذا المقال

مدرس الدكتور محمد لوتي (2026). نماذج اللغة الكبيرة: دمج الخبرة يحسن التشخيص النفسي بدقة أعلى. عرب  
سايكولوجي. تم الاسترجاع من <https://arabpsychology.com/?p=118618>

تخيل أنك طبيب نفسي تواجه حالة معقدة. أمامك وصف تفصيلي لمريض، لكن الوصول إلى التشخيص الدقيق يتطلب خبرة سنوات في تحليل الأعراض، واستبعاد الاحتمالات، وفهم الفروق الدقيقة بين الاضطرابات المختلفة. الآن، تخيل أن لديك مساعداً رقمياً – نموذج لغوي كبير (LLM) - قادراً على معالجة كم هائل من المعلومات الطبية، ولكن هل يمكن لهذا المساعد أن يضاهي دقة التشخيص البشري؟

## منهجية البحث

هذا هو السؤال الذي سعى الباحثون، بقيادة ك.ف. سارما، إلى الإجابة عليه. قام فريق البحث بتقييم أداء النماذج اللغوية الكبيرة الشائعة في مجال التشخيص النفسي. استخدموا سيناريوهات حالات سريرية (وصف تفصيلي لحالات افتراضية للمرضى) مع تشخيصات مرتبطة بها، مأخوذة من كتاب "الحالات السريرية في الدليل التشخيصي والإحصائي للاضطرابات النفسية، الإصدار الخامس المُنقح" (DSM-5-TR Clinical Cases).

ولكن بدلاً من الاعتماد على قدرة النماذج اللغوية على التشخيص بشكل مباشر، قام الباحثون بتجربة طريقة مبتكرة: دمج "أشجار القرارات التشخيصية" (Diagnostic decision trees). هذه الأشجار، المستمدة من "دليل التشخيص التفريقي في DSM-5-TR Handbook of Differential Diagnosis"، تمثل عملية تفكير الطبيب النفسي خطوة بخطوة، حيث يتم استبعاد الاحتمالات بناءً على الأعراض التي تظهر على المريض. قام الباحثون بتعديل هذه الأشجار لتكون مناسبة للاستخدام مع النماذج اللغوية.

تم اختبار ثلاثة نماذج لغوية كبيرة مختلفة. في البداية، تم توجيه النماذج مباشرة لتقديم تشخيصات محتملة لكل حالة سريرية. ثم، تم تزويد النماذج بأشجار القرارات التشخيصية، وتمت مطالبتها باستخدامها للوصول إلى التشخيص. بعد ذلك، قارن الباحثون التشخيصات التي قدمتها النماذج بالتشخيصات الصحيحة، مستخدمين مقاييس إحصائية مثل القيمة التنبؤية الإيجابية (PPV) - وهي نسبة التشخيصات الإيجابية التي تكون صحيحة بالفعل - والحساسية (Sensitivity) - وهي قدرة النموذج على تحديد الحالات الإيجابية بشكل صحيح - وإحصائية F1، وهي مقياس يجمع بين الدقة والحساسية.

## النتائج

أظهرت النتائج أن النماذج اللغوية، عند توجيهها مباشرة لتقديم التشخيصات، حققت حساسية جيدة (76.7٪ للنموذج الأفضل أداءً، وهو gpt-4o)، ولكنها عانت من "التشخيص الزائد" (overdiagnosis) - أي أنها قدمت تشخيصات إيجابية خاطئة بنسبة كبيرة (القيمة التنبؤية الإيجابية 40.4٪ فقط). بمعنى آخر، كانت النماذج تميل إلى رؤية اضطرابات نفسية حتى في الحالات التي لا تستدعي ذلك.

ولكن، عندما تم تزويد النماذج بأشجار القرارات التشخيصية، تحسنت القيمة التنبؤية الإيجابية بشكل ملحوظ (إلى 65.3٪) دون انخفاض كبير في الحساسية (70.9٪). بشكل عام، أظهرت التجارب أن استخدام أشجار القرارات التشخيصية أدى إلى زيادة كبيرة في القيمة التنبؤية الإيجابية، وزيادة في إحصائية F1 في معظم الحالات، وانخفاض طفيف في الحساسية في بعض الحالات.

## دلالات البحث

تشير هذه النتائج إلى أن دمج المعرفة المتخصصة - ممثلة في أشجار القرارات التشخيصية - في عملية التشخيص التي تقوم بها النماذج اللغوية يمكن أن يحسن بشكل كبير من أدائها. فبدلاً من الاعتماد على قدرة النموذج على "التخمين" بناءً

على البيانات التي تدرب عليها، يتم توجيهه من خلال عملية تفكير منظمة ومبنية على الخبرة السريرية.

هذا التحسين يرجع بشكل أساسي إلى قدرة أشجار القرارات على "كبح" التشخيص الزائد، وهو مشكلة رئيسية في النماذج اللغوية عندما يتم استخدامها بشكل مباشر في التشخيص النفسي. على الرغم من أن الحساسية قد انخفضت بشكل طفيف في بعض الحالات، إلا أن الزيادة في القيمة التنبؤية الإيجابية تعني أن النماذج أصبحت أكثر موثوقية في تحديد الحالات التي تتطلب تدخلاً حقيقياً.

هذا البحث يفتح الباب أمام تطوير أدوات قائمة على النماذج اللغوية يمكن أن تساعد الأطباء النفسيين في عملهم، ولكن مع التأكيد على أن هذه الأدوات يجب أن تستخدم كأداة مساعدة وليست كبديل للحكم السريري البشري. فالنماذج اللغوية، حتى مع دمج المعرفة المتخصصة، لا تزال بحاجة إلى إشراف بشري لضمان دقة التشخيص وتقديم الرعاية المناسبة للمرضى.

## Reference

Sarma, K.V. (2026). *Integrating expert knowledge into large language models improves performance for psychiatric reasoning and diagnosis*. Psychiatry Research

DOI: [10.1016/j.psychres.2025.116844](https://doi.org/10.1016/j.psychres.2025.116844)