JALSAT
Journal of Arabic Language Studies and Teaching
Volume: 5, Nomor: 2 (2025)
DOI: 10.15642/jalsat.2025.5.2.234-250

A Comparison of the Accuracy of Generative Artificial Intelligence Models in *Taṣrīf* and the Explanation of *Wazan* Meanings: A Study on Their Application in Arabic Morphology (*Ṣarf*)

مقارنة دقة النموذج للذكاء الإصطناعي التوليدي في التصريف وشرح معاني الأوزان: دراسة توظيفها في علم الصرف

Salman Rizqan Karima, UIN Syarif Hidayatullah, Indonesia
Toto Edidarmo, UIN Syarif Hidayatullah, Indonesia
Raswan, UIN Syarif Hidayatullah, Indonesia

Article Info:

Article History:
Received: 6 Sep 2025
Revised: 22 Oct 2025
Accepted: 21 Nov 2025
Published: 27 Nov 2025

***Corresponding author:**
Name:Salman Rizqan Karima
Email:privatesrk@gmail.com

Abstract

This study analyzes and compares the accuracy of two generative artificial intelligence models, ChatGPT and Deepseek, in producing *taṣrīf* (morphological conjugation) and explaining the meanings of *wazan* (morphological patterns) in Arabic morphology. Employing a descriptive-qualitative method with a content analysis approach, the research uses 30 *fiʿl thulāthī mazīd* (triliteral verbs with augmentation) representing all categories *ṣaḥīḥ, maḥmūz, muḍāʿaf, muʿtal, miṯāl, ajwaf, nāqiṣ*, and *lafīf*. Each verb was tested on *ChatGPT* 4.0 and *Deepseek* V3.1, and the outputs were verified against classical references, including *al-Binā', al-Maqṣūd*, and *al-Amthilah al-Taṣrīfiyyah*. The findings show that both models produce fairly accurate results with distinctive tendencies: Deepseek aligns more closely with classical morphological rules, while ChatGPT provides more diverse and contextual explanations. Pedagogically, both have strong potential to support Arabic morphology learning through interactive and varied outputs, though teacher supervision remains essential to maintain adherence to classical standards.

Keywords: Chat Generative; Artificial Intelligence; Morphology; Content analysis.

This work is licensed under a Creative Commons Attribution 4.0 International License.

المقدمة

إن الذكاء الاصطناعي يلعب دورًا مهمًا في خدمة اللغة العربية عبر التعلم الآلي كتابيًا أو صوتيًا خصوصًا في علم الصرف الذي يتميّز بالتعقيد (حسانين, ٢٠٢٤) . يُعدّ كلٌّ من *ChatGPT* و *Deepseek* من أشهر تطبيقات الذكاء الاصطناعي استخدامًا حاليًا، حيث تجاوز عدد مرات تحميلهما ١٥٠ مليون مرة. يرى بعض الخبراء أن *ChatGPT* يمتلك قدرة واعدة في دعم تعلم اللغة العربية للأغراض الأكاديمية والمهنية، رغم أنه لا يزال بحاجة إلى إشراف وتحقق من قبل المختصين (Yuslina Mohamedi et al., 2023) . كما تشير دراسة حديثة إلى أن *Deepseek* يحظى بتقدير بسبب ذكائه وسهولة الوصول المجاني إليه (Albuhairy & Algaraady, 2025b) .

ChatGPT و*Deepseek* هما نموذجان للذكاء الاصطناعي (AI) يعتمدان على معالجة اللغة الطبيعية (NLP) كأساس لعملهما. التعليم مناسب لتقنيات الذكاء الاصطناعي لأن أنشطته المعرفية يمكن دعمها بقدرات الذكاء الاصطناعي الخوارزمية والمعرفية (Wang et al., 2024, p. 2). كما أثبتت الدردشة(chatbot) فعاليتها في تعزيز المشاركة النشطة للطلاب في تعلم اللغات الأجنبية عبر الإنترنت (Mike Levy & Glenn Stockwell, 2006, p. 253). يُسهّل التعليم الإلكتروني، الذي يتمثل في هذا السياق في الذكاء الاصطناعي، الوصول إلى التعلم المرن والتفاعلي والجذّاب، ويعزّز في الوقت نفسه مشاركة الطلاب، ومهارات القرن الحادي والعشرين، وفعالية المعلم (رمزي أحمد عبد الحي, ٢٠١٠ , p. 133). كما يوفّر التعليم عن بُعد فرصًا مرنة ومتكافئة للتعلّم لجميع الفئات، دون قيود الزمان والمكان (طارق عبد الرؤوف عامر, ٢٠١٨ p. 692 ,).

إنَّ الصرف والتصريف يُعَدّان من أهمّ فروع اللغة العربية، إذ يشكّلان الأساس لفهم بنية الكلمة ومعانيها المتنوّعة. ويُستعمل المصطلحان أحيانًا بمعنى واحد، غير أنّ بينهما فرقًا من حيث الاشتقاق والوظيفة؛ فالصرف علمٌ نظريّ يبحث في أصول بناء الكلمة وقواعدها، بينما التصريف جانبٌ عمليّ يُعنى بتطبيق تلك القواعد في تحويل الكلمة إلى أبنية مختلفة لضروب من المعاني، كتحويل المصدر إلى صيغ الماضي والمضارع والأمر واسم المفعول وغيرهما، وكالنسبة والتصغير(محمد بن مالك الطائي النحوي, ٢٠٠٩ p. 62 ,). ومن ثَمّ، فالعلاقة بينهما علاقة تكامل، إذ لا غنى لأحدهما عن الآخر. وقد ذهب بعض القدماء، كسيبويه، إلى أنّ الصرف هو الجانب القاعديّ النظريّ، والتصريف هو الجانب التطبيقيّ التدريبيّ (هادي نهر, ٢٠١٠ p. 92 ,). ولهذا يُقسَم علم الصرف في الرؤية الحديثة

إلى صرفٍ علميٍّ يُعنى بالقواعد، وصرفٍ عمليٍّ يُعنى بالتطبيق، وكلاهما ضروريّ لإتقان بنية الكلمة العربية وتحقيق الدقة في التعبير اللغوي.

بحسب نظرية الصرف الكلاسيكية، فإنَّ الميزان الصرفي هو أداة وضعها الصرفيّون لقياس وتحليل بُنية الكلمة وتصرّفاتها بدقّة في اللغة العربية (هادي نهر, ٢٠١٠ p. 17) . يُستَخدَم الميزان الصرفي لمعرفة بنية الكلمة والتمييز بين الحروف الأصلية والزوائد. والنموذج القياسي المستخدم هو "ف"، "ع"، "ل" لأن معظم الكلمات العربية ثلاثية الجذور (عبد الشكور معلم عبد فارح, ٢٠٢١ p. 7). الحرف الأول من أصل الكلمة يسمى فاء الكلمة، والثاني عين الكلمة، والثالث لام الكلمة (محمد فاضل السامرائي, ٢٠١٣ p. 11 ,). فبذلك كله قد حاز بنية ثابتة لا يجوز الخروج عن قاعدتها. يمكن أن تدلّ زيادة الحروف في الكلمة على معنى جديد.. فالكلمة "أَكْرَمَ" مشتقة من الصيغة الأصلية "كَرُمَ" من خلال عملية التصريف باتّباع الوزن "أَفْعَلَ". وهذا التغيّر ينقل المعنى من "كَانَ كَرِيمًا" إلى "جَعَلَهُ كَرِيمًا" أو "قَامَ بِتَكْرِيمِه". كما قال ابن جني "واعلم أن لكل حرف من هذه الحروف (الزيادة) موضعا تكثر فيه زيادته (المعنى)" (ابن جني, ١٩٩٨ p. 161 ,).

على الرغم من أن الذكاء الاصطناعي يمكن أن يساعد في رسم قواعد النحو في اللغة العربية، إلا أن نتائجه ليست دقيقة تمامًا ولا تزال تحتاج إلى تدخل (Favirotus Siyam et al., 2024, p. 558). إلا أنه (*ChatGPT*) لا يزال يواجه تحديات في تحليل البُنى النحوية والدلالية في اللغة العربية (رضوان ضيف, ٢٠٢٤) . مع ازدياد الاعتماد عليه، لا بد من تحقق دقة النتائج وتوافقها مع المصادر العربية الكلاسيكية. دور الكتب التراثية أو "الكتب الصفراء" يتمثل في حفظ سَنَد العلم وتوفير أعلى درجات التوثيق والمصداقية (Mochammad Achsan Auza'i et al., 2023).

تحليل بيانات الذكاء الاصطناعي أو التحقق من صحتها هو في الغالب عمل بشري، رغم أن البرمجيات تلعب دورًا كبيرًا فيه (Joseph Teguh Santoso, 2023, p. 91). تشير الدقّة في سياق الذكاء الاصطناعي إلى مدى توافق تنبؤات النموذج مع البيانات الحقيقية، حيث إنّ التقييم غير الدقيق قد يؤدي إلى ضعف الأداء وحدوث مشكلات في التطبيقات الواقعية. وتكمن أهمية تقييم الدقّة في عدة جوانب، منها مقارنة النماذج لاختيار الأفضل من بين عدة خيارات متاحة، وتجنّب الإفراط في التعلّم (Overfitting) لضمان أن يكون أداء النموذج جيدًا ليس فقط على بيانات التدريب بل أيضًا على البيانات الجديدة، إضافةً إلى تحسين النموذج من خلال ضبط إعداداته تدريجيًا لتحقيق نتائج أفضل بشكل مستمر. التوظيف (Usefulness) هو أيضًا مستوى يعبّر عن

إيمان الفرد بأن استخدام تقنية معينة سيؤدي إلى تحسين أدائه أو إنجازه (Dennis A. Adams et al., 1992). تشمل توظيف الادراك بأن التكنولوجيا تُسهّل أداء المهام، وتُفيد في تحسين الأداء، وتُعزّز الإنتاجية نحو الأفضل (Wynne W. Chin & Peter A. Todd, 1995). توظيف نماذج الذكاء الاصطناعي التوليدي في علم الصرف يعني استخدام قدراتها لإنتاج التصريف وشرح الأوزان تلقائيًا وتفاعليًا وفقًا للكتب الصرفية الكلاسيكية.

بالإضافة إلى ذلك، ومن أجل الإجابة عن جانب توظيف الذكاء الاصطناعي في تعليم علم الصرف، فقد استندت أداة البحث في هذه الدراسة أيضًا إلى نظرية تقييم الوسائط التعليمية لـ هاينيش وآخرين (٢٠٠٢). وتشمل هذه النظرية أربعة جوانب رئيسة تُستخدم في تقويم فعالية الوسائل التعليمية، وهي :وضوح الرسالة، ودقة المحتوى بالنسبة للأهداف التعليمية، ومستوى التفاعلية، وسهولة الاستخدام (Heinich et al., 2002, p. 26) . وقد تم اعتماد هذه الجوانب الأربعة كمؤشرات في تقويم مدى إمكانية الاستفادة التربوية من مخرجات *ChatGPT* و *Deepseek* في سياق تعليم علم الصرف. يستنتج الباحث أن دقّة وصلاحية استخدام نماذج الذكاء الاصطناعي مثل *ChatGPT* و *Deepseek* في علم الصرف تعتمد بشكل كبير على مدى مطابقة نتائجها للقواعد الصرفية الكلاسيكية. فَتقييم الدقّة ضروري لضمان اتساق المخرجات مع المعايير العلمية، بينما تُحدّد الصلاحية البيداغوجية مدى إمكانية استخدام تلك النتائج كوسائط تعليمية فعّالة، تفاعلية، وذات معنى.

ليست هذه الدراسة البحثية هي الوحيدة التي تسلط الضوء على حدود أنظمة الذكاء الاصطناعي. في ترجمة اللغة العربية إلى لغات أخرى وخصوصا إلى اللغة الاندونيسية فقد بينت دراسة (Ruhmadi & Al Farisi, 2023) ان *ChatGPT* ما زال يرتكب أخطاء صرفية عديدة خاصة عند التعامل مع المصطلحات العلمية والدينية التي تتطلب دقة عالية في أنماط الاشتقاق والتصريف وهذه النتائج تشير إلى أن تطور التقنية لا يعني خلوها من التحديات الكبرى في معالجة البنى الصرفية المعقدة في اللغة العربية كما تؤكد على ضرورة تدخل الإنسان ولا سيما خبراء اللغة والمترجمين المحترفين لمراجعة وتصحيح الترجمة حفاظا على جودتها فضلا عن أن مثل هذه الدراسات تفتح المجال لتطوير نماذج أكثر حساسية تجاه الخصوصيات اللغوية للعربية بما يضمن انتاج ترجمات أدق و أكثر موثوقية.

أما في دراسة أخرى أجراها (Aunur Shabur Maajid Amadi & Khizanatul Hikmah, 2025) فقد تناولت تصورات الطلاب حول استعمال تقنيات الذكاء الاصطناعي في تعلم اللغة العربية في الجامعات الإسلامية في إندونيسيا وقد أظهرت النتائج أن الطلاب وجدوا فائدة في استخدام *ChatGPT* لفهم البنية الأساسية للصرف والنحو إلا أنهم اشاروا أيضا إلى أن قدرة النموذج ما زالت محدودة في تفسير المعاني العميقة لبعض الصيغ اللفظية خاصة في سياق المصطلحات الدينية والأعمال الأدبية ولذلك يبقى الدور الفعال للأساتذ ضروريا في توجيه الفهم اللغوي المعقد وضمان دقة الاستيعاب.

وكذلك في دراسة (Fitri, 2025) التي جاءت على شكل مراجعة أدبية منهجية حول دمج تقنيات الذكاء الاصطناعي في تعليم اللغة العربية، بما في ذلك استخدام *ChatGPT* و*Deepseek*، فقد بُيِّن أن توظيف هذه النماذج يمكن أن يسهم في رفع فعالية عملية التعلم، خاصة في الجوانب المتعلقة بالصرف والنحو. غير أن الدراسة نفسها شددت على أن الاعتماد الكلي على هذه التقنيات قد لا يكون كافيًا، إذ تظل الحاجة قائمة إلى التحقق اليدوي من قبل المتخصصين لضمان دقة المعلومات اللغوية الناتجة، وذلك بالنظر إلى احتمالية وقوع أخطاء أو تبسيطات مفرطة في المخرجات التي تقدمها هذه النماذج.

وفي دراسة أُجريت من قِبَل (Albuhairy & Algaraady, 2025a) بعنوان *Deepseekمقابل ChatGPT: الفعالية المقارنة في التحليل الاستدلالي لاكتساب اللغة الثانية لدى البالغين*، تمت المقارنة بين فعالية هذين النموذجين في كشف الأخطاء اللغوية التي يرتكبها متعلمو العربية كلغة ثانية. وقد أظهرت النتائج أن نموذج *Deepseek* يتفوق في تحديد الأخطاء السياقية، مثل نقل ترتيب الكلمات من اللغة الأم وفق نموذجSOV ، في حين برز *ChatGPT* بقدرته الأفضل على تقديم تغذية راجعة تعليمية وذات صلة. ومع ذلك، أوضحت الدراسة أن كلا النموذجين ما زالا بحاجة إلى تعليمات دقيقة (*prompts*) لمعالجة الأخطاء الدلالية والبراغماتية، بما في ذلك استخدام أدوات التعريف بشكل خاطئ أو حالات عدم التوافق بين اللهجات.

لذلك، ونظرًا لعدم وجود دراسة تناولت بشكل خاص جانب الذكاء الاصطناعي في سياق الصرف الاشتقاقي في اللغة العربية باستخدام الكتب الصفراء الكلاسيكية، فإن هذا البحث يهدف بشكل خاص إلى مقارنة نموذجين بارزين من نماذج AI ، وهما *ChatGPT* و*Deepseek*، في إنتاج التصريف = وشرح معاني الأوزان في سياق علم الصرف. إلى جانب التحليل اللغوي، يقدّم هذا البحث

أيضًا شرحًا تربويًا حول كيفية توظيف AI في تعليم علم الصرف، وإن لم يُطبّق ذلك مباشرة في البيئة الصفّية. ومن خلال هذا النهج، يجمع البحث بين الجانب اللغوي، والمعرفة التراثية، والإمكانات التعليمية التي قد يقدّمها نموذج الذكاء الاصطناعي.

منهجية البحث

اعتمد هذا البحث على المنهج الوصفي الكيفي باستخدام أسلوب تحليل المحتوى بهدف تقويم دقة نتائج التصريف ومعاني الأوزان التي ينتجها نموذجا الذكاء الاصطناعي *ChatGPT* و *Deepseek*. وتكوّنت عيّنة البحث من ثلاثين فعلاً ثلاثياً مزيداً اختيرت بطريقة قصدية، وقد اختُبرت الأفعال باستخدام برومبت موحّد وفق أسلوب zero-shot لضمان الحياد والاتساق. وشملت البيانات محلّ التحليل مخرجات النموذجين ومقارنة نتائجها بثلاثة من الكتب الصرفية الكلاسيكية وهي الأمثلة التصريفية والبناء والمقصود.

لضمان الموضوعية وضع نظام تقييم كمي وكيفي يتكون من فئتين رئيسيتين:

أ. للتصريف.

الدرجة	المعيار	الوصف
٤ نطق	مطابق كامل	الناتج صحيح وكامل حسب القواعد/الكتب المعتمدة، ويشتمل على جميع الصيغ المطلوبة.
٣ نطق	مطابق غير كامل	الناتج صحيح في ذاته لكن لا يشتمل على جميع الصيغ.
٢ نطق	جزئي صحيح وجزئي خطأ	يحتوي الناتج على مزيج بين الصواب والخطأ.
١ نطق	غير مطابق/كله خطأ	الناتج كله خاطئ أو غير موافق للقواعد/الكتب المعتمدة.

ب. للمعنى الوزن

الدرجة	المعيار	الوصف
٣ نطق	مطابق	الجواب صحيحًا وموافقًا للقواعد أو للكتب المعتمدة في الصرف، سواء اقتصر على معنى واحد إذا لم يكن للوزن إلا ذلك المعنى، أو جمع أكثر من معنى إذا كان للوزن تعدد في الدلالات.

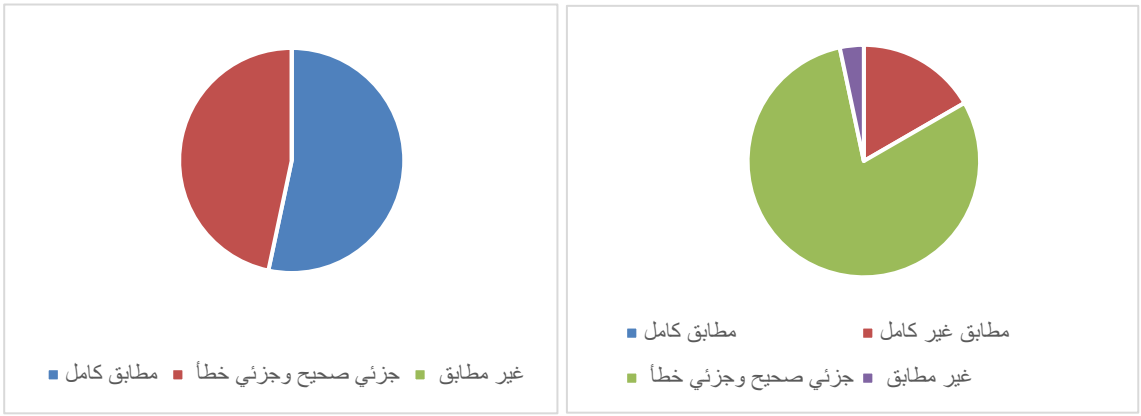
٢ نطق	جزئي صحيح وجزئي خطأ	اشتملت الإجابة على مزيج بين الصواب والخطأ، بحيث يُذكر بعض المعاني الصحيحة مع أخرى غير صحيحة.
١ نطق	غير مطابق/كله خطأ	كانت الإجابة خاطئة تمامًا أو غير موافقة للقواعد والكتب المعتمدة في علم الصرف، كإيراد صيغ أو معانٍ لا وجود لها في التراث الصرفي الكلاسيكي.

جُمعت البيانات بوساطة الملاحظة والتوثيق وتحليل الوثائق، ثم حُللت وفق نموذج مايلز وهوبرمان الذي يتضمن تخفيض البيانات وعرضها والتحقق من الاستنتاجات. ودعمت مصداقية النتائج من خلال التثليث بالخبراء بمشاركة ثلاثة من المتخصصين في علم الصرف. وتمثلت أداة البحث الرئيسة في الباحث نفسه، بالإضافة إلى استمارات تقييم دقة التصريف ومعاني الأوزان المستندة إلى المراجع الصرفية المعتمدة.

النتائج والمناقشة

ننتقل الآن إلى عرض نتائج البحث ومناقشتها وفي هذا السياق يقدم فيما يلي الصور التي توضح مقارنة الأوزان المختبرة مع عددها وكيفية معالجة كل من *Deepseek* و *ChatGPT* لها.

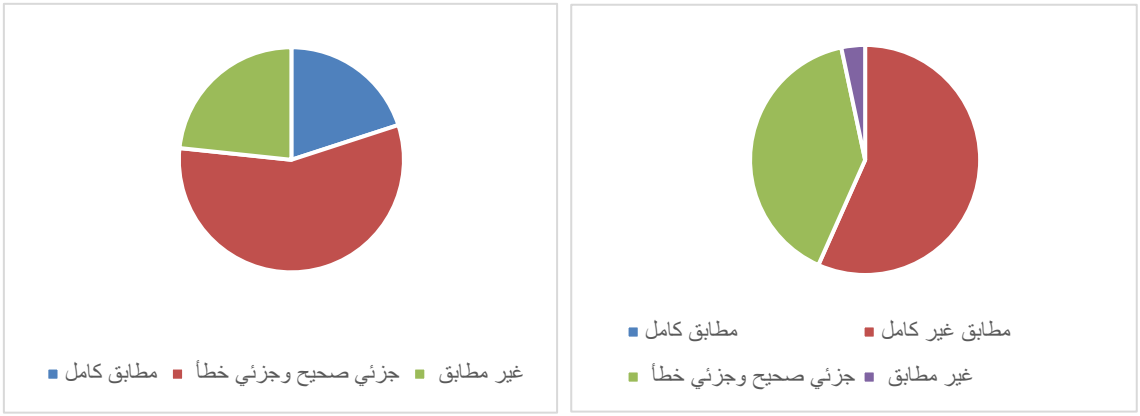
الصورة ١: مخرجات *ChatGPT* في الصريف وشرح معاني الأوزان



أظهرت نتائج التحليل أن قدرة *ChatGPT* في إجراء التصريف ما زالت منخفضة إذ حصل على أربع وخمسين نقطة من مئة وعشرين، وكانت أغلب الإجابات صحيحة جزئياً ولم تُسجَّل أي إجابة صحيحة تماماً. وتنوّعت الأخطاء في المصدر الميمي واسم المكان واسم الزمان وفي تطبيق قواعد العلة.

أما في جانب شرح معاني الأوزان الصرفية فقد كان الأداء أفضل إذ نال ستاً وسبعين نقطة من تسعين، وجاءت أكثر الإجابات صحيحة موافقة للقواعد مع وجود بعض الأخطاء الطفيفة. وبذلك يتبيّن أن جودة *ChatGPT* في فهم المعاني أعلى منها في أداء التصريف.

الصورة ٢: مخرجات *ChatGPT* في الصريف وشرح معاني الأوزان



أظهرت نتائج البحث أن قدرة *Deepseek* في إجراء التصريف ما زالت منخفضة إذ حصل على إحدى وستين نقطة من مئة وعشرين وكانت أغلب الإجابات صحيحة جزئياً ولم تسجل أي إجابة صحيحة تماماً. وظهرت بعض الأخطاء في صيغ الأفعال المزيدة وبعض المشتقات.

وفي جانب شرح معاني الأوزان الصرفية جاء الأداء ضعيفاً نسبياً إذ نال تسعاً وخمسين نقطة من تسعين وكانت الإجابات الصحيحة التامة أقل من الخاطئة أو الجزئية. وبذلك يتبيّن أن *Deepseek* لم يستوعب بعد معاني الأوزان على وفق ما ورد في كتب الصرف الكلاسيكية وأن جودة مخرجاته ما زالت محدودة.

التحليل عن مقارنة دقّة كلا النموذجين

المحور	***ChatGPT***	***Deepseek***
التصريف	٩٠/٧٦	١٢٠/٥٤
شرح معنى الوزن	٩٠/٥٩	١٢٠/٧٦

بناءً على نتائج البحث يمكن القول إنّ ديبسيك كان متفوقًا في جانب التصريف إذ حصل على إحدى وستين درجة من مئةٍ وعشرين، مقابل أربعةٍ وخمسين درجة فقط لـ *ChatGPT*. وأمّا في جانب شرح معاني الأوزان فقد كان *ChatGPT* متفوقًا إذ حصل على ستٍّ وسبعين درجة من تسعين،

بينما لم يحصل *Deepseek* إلا على تسعٍ وخمسين درجة. وهذا يدل على أنّ كلا النظامين لهما ميزات في جانب خاص، وبذلك يُكمل أحدهما الآخر في مجال علم الصرف.

وبحسب البيانات فإن *Deepseek* أكثر تفوقا في جانب التصريف حيث حصل على ٦١ من ١٢٠ مقارنة ب *ChatGPT* الذي حصل فقط على ٥٤ من ١٢٠ وهذا يدل على أن *Deepseek* أقدر نسبيا في إنتاج الأبنية الصرفية الأساسية بشكل أكثر اتساقا رغم أنه ما زال يرتكب كثيرا من الأخطاء في الفئات الخاصة مثل المصدر الميمي واسم المكان واسم الزمان وعلى العكس من ذلك فإن *ChatGPT* أكثر تفوقا في جانب شرح معنى الوزن حيث حصل على ٧٦ من ٩٠ متقدما كثيرا على *Deepseek* الذي لم يحصل إلا على ٥٩ من ٩٠ فقد كان *ChatGPT* أكثر ميلا لتقديم تعريفات مطابقة للتراث الكلاسيكي رغم أنها أحيانا تكون مصحوبة بمصطلحات معاصرة أو تفسيرات إضافية بينما *Deepseek* كان يميل إلى المصطلحات المعاصرة.

مزايا نظامي الذكاء الاصطناعي *ChatGPT* و *Deepseek* في سياق هذا البحث على النحو: الآتي:

١. **السرعة في إعطاء الجواب**

يقدمان الجواب في ثوان قليلة، مما يسهّل على المستخدم الحصول على النتيجة من غير معالجة يدوية طويلة.

٢. **سهولة الوصول والتوفّر**

كلا النظامين متاح على الشبكة ويمكن استعمالهما مجانًا، فيسهل استخدامهما في السياقات الأكاديمية وغير الأكاديمية.

٣. **دقّة نسبيّة في الأوزان العامّة**

مع أنّ الأخطاء موجودة، إلا أنّ النظامين متسقان في التصريف للأوزان الصحيحة والأشكال المشهورة، فيمكن الاعتماد عليهما أداةً أولية في التحليل الصرفي.

٤. **إمكانات في التعليم والبحث**

يمكن توظيفهما وسيلة مساعدة في تعليم الصرف، خاصة في التمرينات الأساسية والمحاكاة، وكذلك مادة للمقارنة في الدراسات اللغوية.

وأما أنماط الأخطاء العامة في النظامين فهي على النحو الآتي:

١. **أخطاء في القواعد الصرفية**

وهي عجز النظام عن معالجة الإبدال والإدغام والتغييرات الخاصة في الفعل المعتل (الإعلال). كما تظهر الأخطاء في تكوين فعل الأمر والنهي، ولا سيما في تطبيق حذف حرف المضارعة وتحديد الحركة الأخيرة. بالإضافة إلى ذلك فإن النظام محدود في معالجة اجتماع حروف العلة في الفعل المزيد، خصوصًا عند وجود علتين معًا.

٢. أخطاء في الحركات والأبنية

وتظهر في عدم الاتساق في ضبط الحركات، مثل ظهور الفتحة في موضع يجب فيه الضمة. كما وُجدت أخطاء في تكوين اسم الزمان واسم المكان والمصادر، إذ قد لا يظهر الشكل أصلًا، أو يظهر ولكنه لا يوافق الوزن القياسي.

٣. أخطاء في معنى الأوزان

فكثيرًا ما يقدّم النظام شروحًا لا توافق المصطلحات الصرفية الكلاسيكية، مثل ترجمة معنى المشاركة بلفظ غريب أو فهم المطاوعة بمعنى "علاقة". وأحيانًا يكون التعريف صحيحًا والمصطلح خطأ، أو يكون المصطلح صحيحًا ولكن الشرح منحرفًا عن التقليد الصرفي واستعمال مصطلحات غير كلاسيكية.

٤. أخطاء منهجية وعدم الاتساق

إذ يميل النظام إلى التعميم المفرط بإخراج صيغة واحدة للتصريف من غير مراعاة تنوّع القواعد الصحيحة. كما تظهر عدم الاتساق بين جداول التصريف والشروح المرافقة لها، مما يقلّل من موثوقية النتائج.

أظهرت نتائج *ChatGPT* تفوقا كميا في شرح معاني الأوزان مقارنة ب *Deepseek* حيث سجل عددا أكبر من الإجابات المتوافقة مع كتب الصرف الكلاسيكية ومع ذلك لا يزال النموذجان يعانيان من نقص في الدقة خاصة عند التعامل مع الأوزان المركبة أو الأقل شيوعا كثير من الشروحات التي قدمها النموذجان تستند إلى السياق المعنوي العام ولا تعكس المعنى الاشتقاقي الحقيقي المرتبط بالبنية الصرفية للوزن وهذا يدل على أن الذكاء الاصطناعي يعتمد في الغالب على استخدامات معجمية حديثة او استعمالات لغوية في سياقات عامة وليس على المفهوم الصرفي الدقيق الذي تقره المراجع التراثية.

كما أشرنا سابقا تبين أن نموذج *Deepseek* يظهر أداء أعلى مقارنة ب *ChatGPT* من حيث دقة التصريف الصرفي بناء على نوع الوزن وتبرز أفضلية *Deepseek* بشكل خاص في الأوزان القياسية مثل فاعل وأفعل وهي الأوزان التي تظهر كثيرا في تعليم اللغة العربية الرسمي وقد تمكن

Deepseek من تحقيق مستوى أعلى من الدقة في أنتاج الصيغ المختلفة مثل الفعل الماضي والمضارع والمصدر واسم الفاعل واسم المفعول بشكل منتظم ومتسق.

ويمكن تفسير هذا التميز علميا من خلال عدة عوامل اولا يعتمد *Deepseek* على بنية الخبراء المتعددين Mixture of Experts والانتباه الكامن المتعدد الرؤوس Multi head Latent Attention مما يتيح معالجة أكثر تركيزا وكفاءة للأنماط الصرفية ثانيا تم تدريب النموذج باستخدام ١٤,٨ تريليون وحدة لغوية عالية الجودة من مصادر متنوعة بما في ذلك نصوص عربية رسمية مما يعزز فهمه للأوزان الصرفية وأنماطها ثالثا أن نجاح *Deepseek* في مجال التصريف مدعوم أيضا بتقنيات التقطير القائم على الاستدلال reasoning based distillation وسلاسل التفكير Chain of Thought CoT التي تسهم في تعزيز الاتساق في توليد البنى الصرفية .وقد بينت أبحاث مثل دراسة (Hasanaath et al., 2025) أن *Deepseek* يتفوق كذلك في المهام اللغوية المتعلقة باللغة العربية مثل تصنيف أقسام الكلام part of speech tagging وتفسير المعاني المتعددة للكلمة word sense disambiguation حيث حقق نتائج دقيقة أعلى بشكل ملحوظ مقارنة بنماذج أخرى مثل GPT 4o او LLaMA وبناء على ذلك فان تفوق *Deepseek* في توليد الاشكال الصرفية في اللغة العربية ليس أمرا عشوائيا بل هو نتيجة لتصميم معماري مدروس واستراتيجية تدريب ملائمة لمعالجة اللغة العربية.

يعزى أحد أسباب تفوق نموذج *Deepseek* إلى نوع ومصدر البيانات التي تم تدريبه عليها يشير (Mona Elswah, 2024) من كلية كينيدي بجامعة هارفارد إلى أن نماذج اللغة العربية التي يتم تدريبها فقط على بيانات وسائل التواصل الاجتماعي تميل إلى فقدان القدرة على التقاط انماط اللغة الكلاسيكية ولذلك فإن النماذج مثل *ChatGPT* التي لم تدرب بشكل خاص على بيانات إسلامية أو لغوية عربية تراثية غالبا ما تنتج صيغ غير مستقرة أو غير فصيحة لا سيما في المشتقات مثل اسم المكان واسم الزمان كما أن النتائج التي ناقشها الباحث تتماشى مع الدراسات السابقة التي تم الاشارة إليها ومنها دراسة (Albuhairy & Algaraady, 2025a) في مقالتهما *Deepseek* vs *ChatGPT* حيث أكدا أن *Deepseek* يتفوق في اكتشاف الاخطاء الصرفية والنحوية الشائعة لدى المتعلمين غير الناطقين بالعربية خاصة في سياق تعليم اللغة العربية كلغة ثانية وقد صرحا بأن *Deepseek* أكثر قدرة على كشف أخطاء النقل النحوي الناتجة عن اللغة الأم L1 induced syntactic transfer errors مقارنة بـ *ChatGPT* .

تقييم إمكانات توظيف الذكاء الاصطناعي في تعليم علم الصرف

في هذا القسم تم تحليل أداء نموذجي الذكاء الاصطناعي *ChatGPT* و *Deepseek* في سياق تعليم علم الصرف وذلك من خلال أربعة جوانب رئيسية تمثل معايير التقييم التربوي والتقني. يهدف هذا التحليل إلى إظهار مواطن القوة والضعف في كل نموذج وتحديد مدى صلاحيتهما لدعم التعلم الذاتي وتحقيق الاهداف التعليمية. ومن حيث التفاعلية فإن كلا النموذجين قادر على فهم سياق السؤال وتقديم أجوبة واضحة لكن *ChatGPT* يتفوق بقدرته على بناء حوار تفاعلي واستباق حاجات المتعلم بينما يميل *Deepseek* إلى الأسلوب الأحادي الأقل ديناميكية.

من حيث ثبات النتائج فان *ChatGPT* قد ينتج صيغ صرفية متعددة عند تكرار الطلب مما يسبب ارباكا للمتعلمين المبتدئين لكنه في شرح معاني الاوزان يقدم اجوبة أكثر دقة وأقرب إلى المراجع الكلاسيكية. أما *Deepseek* فيتميز بالثبات في عرض الصيغ الصرفية مع وجود بعض الأخطاء البسيطة لكنه في تفسير الأوزان كثيرا ما يقدم معاني لا توافق المصادر التراثية. وهذا يعني أن لكل نموذج ميزة ونقصا في جانب الثبات والدقة.

وأما من حيث ملاءمة التعلم الذاتي وسهولة تنفيذ المهام فكلا النموذجين يساعدان المتعلم على فهم المبادئ الأساسية بسرعة وبوضوح ويقدمان شروحا عملية ومباشرة. كما يتيح *ChatGPT* خيارات عرض متنوعة مثل الجداول او الملفات مما يسهل التوثيق والافادة. غير أن كلاهما ما زال عاجزا عن معالجة موضوعات صرفية متقدمة مثل اسم الزمان واسم المكان والمصدر الميمي بشكل دقيق. ولذلك فهما صالحان للمستويات الأولى لكن يحتاج المتعلم إلى مراجعة النتائج بالعودة إلى المصادر الأصيلة أو بمتابعة من المعلم عند التعمق في المسائل الصرفية.

يتميز *ChatGPT* من حيث التفاعلية إذ يستطيع بناء حوار ديناميكي مع المستخدم وتقديم المساعدة الإضافية بشكل آني وتكييف ردوده بناء على احتياجات وسياق المتعلم. ويتماشى هذا مع نتائج (Malik & Dahiya, 2024) التي اشارت الى ان *ChatGPT* يوفر "تجربة روبوت محادثة متقدمة" تدعم عملية التعليم والتعلم وتعزز من تفاعل الطلاب في البيئة الأكاديمية. بالإضافة إلى ذلك تؤكد مراجعة نشرت فيFrontiers in Education من دراسة (Bettayeb et al., 2024) ان دمج *ChatGPT* في التعليم يحقق فوائد متعددة مثل تقديم الدعم المخصص وزيادة التفاعل وتجربة تعلم أكثر تكيفا.

يعد كلا النموذجين *ChatGPT* و *Deepseek* مفيدين جدا في سياق التعلم الذاتي للمفاهيم الاساسية في اللغة العربية مثل الحركات الاعرابية المرفوع وتصريف الافعال الاساسية وفهم الاوزان الصرفية الشائعة كما يدعمان تنفيذ المهام التعليمية من خلال الردود السريعة وسهولة الوصول

والمرونة في الاستخدام مما يساعد المتعلمين المبتدئين بشكل كبير. فالذكاء الاصطناعي مناسب لدعم المراحل الاولى من التعلم لكنه لا يكفي لمعالجة المستويات المتقدمة دون مرافقة المعلم والرجوع إلى المصادر التراثية. وبذلك يمكن ل *ChatGPT* و *Deepseek* أن يلعبا دورا فعالا كادوات مساعدة تعليمية إذا وضعا في إطار تعليمي مناسب. ومع ذلك فان نتائجهما في الدراسات الاكاديمية المتقدمة أو المناقشات اللغوية العليا ينبغي أن تتحقق على يد الخبراء لضمان الدقة والمصداقية العلمية.

يمكن الاستفادة من *ChatGPT* و *Deepseek* في تعليم الصرف، ليس فقط في تدريب الطلاب على تصريف الأفعال، بل أيضًا في فهم المعاني الدلالية للأوزان المختلفة. يتميّز *ChatGPT* بالتفاعلية حيث يتيح للطالب طرح أسئلة مباشرة حول وظائف الأوزان مثل فاعل، تفاعل، استفعل، والحصول على شروح دلالية مصحوبة بأمثلة. أما *Deepseek* فيبرز من حيث الثبات والاتساق في إنتاج الأشكال الصرفية المنتظمة، مما يجعله أداة مناسبة للتحقق من صحة التصريف والمقارنة مع المصادر التراثية مثل البناء والمقصود. كما يمكن توظيفهما في التدريبات التلقائية، وتحليل الأخطاء، وتنمية مهارة التحليل اللغوي لدى الطلاب من خلال المناقشات الصفية التفاعلية.

تشير نتائج البحث إلى أن دمج أدوات الذكاء الاصطناعي مثل *ChatGPT* و *Deepseek* يسهم في تسريع الوصول إلى المعلومة، وتوفير تدريبات تكيفية، وزيادة مرونة التعلّم، خاصة لفئة المبتدئين. غير أنّ النموذجين يظلان محدودين عند معالجة القضايا اللغوية المتقدمة، مما يستلزم وجود إشراف المعلّم والرجوع إلى المراجع التراثية لضمان الدقة وعمق الفهم. وبذلك يمكن اعتبار *ChatGPT* و *Deepseek* أدوات تعليمية فعّالة إذا استُخدِمت ضمن إطار تربوي متوازن يجمع بين التقنية الحديثة والإشراف البشري.

الخاتمة

بناء على نتائج البحث حول دقة التصريف وشرح معاني الأوزان التي ينتجها كل من *ChatGPT* و*Deepseek* ، يمكن القول إنّ مخرجاتهما تتميّز بدرجة مقبولة من الدقة، وإن كان هناك اختلاف في عمق التحليل واختيار المصطلحات. فقد ظهر أن *Deepseek* أكثر التزامًا بالقواعد الصرفية الكلاسيكية، بينما يميل *ChatGPT* إلى تقديم شروح سياقية أكثر تنوعًا. وبصورة عامة، فإن كلا النموذجين يقدّمان أدوات مساعدة مهمّة للمتعلمين، غير أنّ الاعتماد الكامل عليهما دون مراجعة المعلم أو المراجع التراثية قد يوقع الطالب في أخطاء تؤثر في فهمه الأساسي.

أما في المجال التعليمي، فإن استخدام هذين النموذجين يفتح آفاقًا جديدة في تدريس علم الصرف، خصوصًا في مجال تدريبات التصريف وصياغة الأسئلة والمناقشات التفاعلية. يمكن للمدرس الاستفادة منهما لتسهيل التعلم وتوفير وقت الإعداد، شريطة أن يتحقق دائمًا من صحة النتائج بالرجوع إلى الكتب الصرفية المعتمدة مثل *المقصود* و*البناء* و*الأمثلة*. كما يمكن للطلاب توظيفهما كوسيلة تدريب ذاتي لتعميق مهاراتهم الصرفية وفهم الوظائف الدلالية للأوزان، مع الوعي بضرورة المراجعة والتثبت من صحة المعلومات.

وأما على صعيد تطوير الذكاء الاصطناعي، فإن من المهم تعزيز بيانات التدريب بالنصوص الصرفية الكلاسيكية لتصبح المخرجات أكثر دقة في معالجة الأفعال المعتلة والأوزان المركبة. كما يُستحسن ابتكار آليات تحقق داخلية تحدّ من الأخطاء وتزيد من موثوقية النتائج. ومن جهة أخرى، فإن الباحثين القادمين مدعوون إلى توسيع نطاق الدراسة بمقارنة نماذج أخرى أو اختبار فاعليتها في مجالات لغوية مختلفة مثل النحو والبلاغة والدلالة، بالإضافة إلى إجراء بحوث تطبيقية في الصفوف الدراسية لمعرفة أثر هذه النماذج مباشرة في تحسين مستوى الطلاب..

المراجع

Albuhairy, M. M., & Algaraady, J. (2025a). *Deepseek* vs. *ChatGPT*: Comparative Efficacy in Reasoning for Adults' Second Language Acquisition Analysis. *مجلة العلوم التربوية و الدراسات الإنسانية*, ٤٤, ٨٦٤–٨٨٣. https://doi.org/10.55074/hesj.vi44.1313

Albuhairy, M. M., & Algaraady, J. (2025b). *Understanding User Perceptions of Deepseek: A Mixed-Methods Sentiment and Thematic Analysis*. https://doi.org/10.36227/techrxiv.174140724.41500096/v1

Aunur Shabur Maajid Amadi & Khizanatul Hikmah. (2025). Persepsi Mahasiswa Tentang Pemanfaatan Teknologi AI dalam Pembelajaran Bahasa Arab di Perguruan Tinggi Islam Indonesia. *Journal of Education Research*, *6*(2), 292–301. https://doi.org/10.37985/jer.v6i2.2343

Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024). Exploring the impact of *ChatGPT*: Conversational AI in education. *Frontiers in Education*, *9*, 1–15. https://doi.org/10.3389/feduc.2024.1379796

Dennis A. Adams, R. Ryan Nelson, & Peter A. Todd. (1992). Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication. *MIS Quarterly*, *16*(2), 227–247.

Favirotus Siyam, F., Hidayat, R., Rochmat, C. S., Maulaya, R. D., Avilya, A., & Maulidi, M. B. (2024). Accuracy Analysis of Artificial Intelligence in Arabic Language Translation and Grammatical Rule Mapping. *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, *16*(2), 558. https://doi.org/10.24042/albayan.v16i2.24588

Fitri, A. T. (2025). *Tinjauan Literatur: Integrasi Kecerdasan Buatan (AI) Dalam Pembelajaran Bahasa Arab*. *9*(1).

Hasanaath, A., Alansari, A., Ashraf, A., Salmane, C., Luqman, H., & Ezzini, S. (2025). *AraReasoner: Evaluating Reasoning-Based LLMs for Arabic NLP* (No. arXiv:2506.08768). arXiv. https://doi.org/10.48550/arXiv.2506.08768

Heinich, Robert, Molenda, Michael, Russell, James D, Smaldino, & Sharon E. (2002). *Instructional Media and Technologies for Learning* (7th edition). Merrill / Prentice Hall.

Joseph Teguh Santoso (with Mars Caroline Wibowo). (2023). *Kecerdasan Buatan (Artificial Intelligence)* (Muhammad Sholikan, Ed.). Yayasan Prima Agus Teknik bekerja sama dengan Universitas STEKOM.

Malik, N., & Dahiya, M. (2024). Interacting with chat GPT: Impact on students and teaching-learning. *Sigma Journal of Engineering and Natural Sciences – Sigma Mühendislik ve Fen Bilimleri Dergisi*, *42*(6), 1736–1748. https://doi.org/10.14744/sigma.2023.00147

Mike Levy & Glenn Stockwell. (2006). *خيارات (CALL): أبعاد تعلم اللغة بمساعدة الحاسب الآلي* (حسيب علي, Trans.). مكتبة الملك فهد الوطنية. *وموضوعات في تعلم اللغة بمساعدة الحاسب الآلي*

Mochammad Achsan Auza'i, Uswatun Khasanah, & Muhammad Luki Faturrokhman. (2023). Pengkajian Kitab Kuning Sebagai Tonggak Keilmuan di Pondok Al Utsmani. *INTEGRATIA: Journal of Education, Human Development, and Community Engagement*, *1*(2), 181–192. https://doi.org/10.71155/integratia.v1i2.45

Mona Elswah. (2024). *Does AI Understand Arabic? Evaluating the Politics Behind the Algorithmic Arabic Content Moderation* (Working Paper Nos. 2024–01; Carr Center Discussion Paper Series). Carr Center for Human Rights Policy, Harvard Kennedy School, Harvard University.

Ruhmadi, A., & Al Farisi, M. Z. (2023). Analisis Kesalahan Morfologi Penerjemahan Arab–Indonesia pada *ChatGPT*. *Aphorisme: Journal of Arabic Language, Literature, and Education*, *4*(1), 55–75. https://doi.org/10.37680/aphorisme.v4i1.3148

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Wynne W. Chin & Peter A. Todd. (1995). On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research: A Note of Caution. *MIS Quarterly*, *192*(2), 237–246.

Yuslina Mohamedi, Zainurrijal Abd Razaki, Sulaiman Ismail, Mohamed Haji Ibrahim, Wan Azura Wan Ahmad, & Zulkipli Md Isa. (2023). *ChatGPT*-AI Technology in Teaching Arabic for Specific Purposes from the Arabic Expert's Perspective. *Al-Qanatir: International Journal of Islamic Studies*, *32*(2), 213–225.

ابن جني. (١٩٩٨). *التصريف الملوكي* جزيع سقال ;.Ed ,الطبعة الأولى .(دار الفكر العربي.

حسانين, ن. ح. ع. (٢٠٢٤). توظيف الذكاء الاصطناعي لخدمة اللغة العربية " من وجهة نظر المعلمين ". *مجلة الناطقين بغير اللغة العربية*, *٧*(٢٠), ١٩١–٢١٢ . https://doi.org/10.21608/jnal.2024.339463

رضوان ضيف. (٢٠٢٤). الذكاء الاصطناعي ومشاكل التشريح اللساني في اللغة العربية : *ChatGPT* أنموذجاً *Djoussour El-Ma'arifa .(جسور المعرفة)*, ١ (٣), ٢٨٢–٢٩٧.

رمزي أحمد عبد الحي. (٢٠١٠). *التعليم عن بعد في الوطن العربي وتحديات القرن الحادي والعشرين*. مكتبة الأنجلو المصرية.

طارق عبد الرؤوف عامر. (٢٠١٨). *التعليم عن بعد والتعليم المفتوح*. دار الوفاء لدنيا الطباعة والنشر.

عبد الشكور معلم عبد فارح. (٢٠٢١). *الصرف الميسر: تقريب لأبنية الأفعال لابن مالك بأسلوب سهل مع الجداول والتدريبات* (الطبعة الثانية). دار التعليم.

محمد بن مالك الطائي النحوي. (٢٠٠٩). *إيجاز التعريف في علم التصريف*. مكتبة الثقافة الدينية.

محمد فاضل السامرائي. (٢٠١٣). *الصرف العربي: أحكام ومعانٍ* (الطبعة الأولى). دار ابن كثير.

هادي نهر. (٢٠١٠). *الصرف الوافي: دراسات وصفية تطبيقية* (الطبعة الأولى). عالم الكتب الحديث.