مجلة جامعة أم القرى لعلوم اللغات وآدابها

الموقع الإلكتروني: https://uqu.edu.sa/jll

معرفة المتسبب في الحدث: تقييم قدرة نموذج الـ (ChatGPT)على إدراك الأدوار المحورية في الجملة العربية

Identifying the Causer of the Event: Evaluating the Performance of ChatGPT on Thematic Roles Recognition

د. يحيى بن علي آل مريع عسيري*

أستاذ مساعد، قسم اللغة العربية وآدابها، كلية الآداب والعلوم الإنسانية، جامعة الملك خالد، المملكة العربية السعودية

Dr. Yahya Ali Al-Murayya Aseri*

Assistant Professor, Department of Arabic Language, College of Arts and Humanities, King Khalid University, Saudi Arabia

تاريخ التقديم: 28/11/2023 تاريخ ارسال التعديلات: 10/2/2024 تاريخ القبول:11/3/2024

Received:28/11/2023 Revised:10/2/2024 Accepted: 11/3/2024

الملخص:

تهدف هذه الدراسة إلى اختبار قدرة نماذج الذكاء الاصطناعي اللغوية الضخمة(LLMS) ، وتحديدًا نموذج الـ(ChatGPT)، على معرفة المتسبب في الحدث دون الحاجة إلى الاستعانة بخوارزميات التحليل النحوي وخوارزميات التحليل الدلالي ودون الاستعانة بالموارد المعجمية المعدّة مسبقًا لهذا الغرض. وقد اختُبِرَتْ قدرة هذا النموذج على ما يقرب من 200 جملة تشتمل على 43 فعلا/مشتقًا تتطلب طبيعتها المعجمية وجود متسبب في الحدث، ومن ثمَّ مقارنة أدائه بنموذج معياري أعطي لاثنين من اللغويين للإجابة عنه. وقد كشفتْ نتائجُ هذه الدراسة عن ضعف قدرة هذا النموذج الاصطناعي على معرفة المتسبب في الحدث بعد معرفة الكلمة الدالة على الحدث، حيث تبين أن هذا النموذج، مقارنةً بأداء الإنسان، قد حقق ما متوسطه (F-Score=0.53) في مهمة تحديد الكلمة الدالة على الحدث، وحقق متوسط (F1-Score=0.63) في معرفة المتسبب في الحدث. وهذا يكشف لنا أن نموذج الـ(ChatGPT)، بوصفه أحد نماذج الذكاء الاصطناعي الضخمة، وإن أظهر قدرة على توليد النصوص تقارب قدرة الإنسان، ما يزال غير قادر على فهم العلاقات الدلالية إدراكها داخل الجملة بصورة تقارب قدرة الإنسان على الفهم والاستنتاج. ومع هذا، تقترح هذه الدراسة الاستفادة من هذا النموذج في تحليل الأدوار المحورية، على أن يكون ذلك بتدخل الإنسان، وذلك بعرض النتائج على متخصص يقوم بمراجعتها وتصحيحها، وهو ما قد يساعد في توسيم البيانات تمهيدًا لاستخدامها في مهام تدريب الآلة .

الكلمات المفتاحية: الحدث، الأدوار المحورية، المتسبب في الحدث، نماذج الذكاء الاصطناعي اللغوية الضخمة، ChatGPT.

Abstract:

This study aims to test the capability of large language models (LLMS), specifically the ChatGPT model, to determine the causer of an event without relying on morphological and syntactic algorithms and lexical resources. The model's ability was evaluated on roughly 200 sentences containing 43 verbs/derivatives of verbs that lexically require a causer in the event structure. The performance of the model was then compared to a gold standard model annotated by two linguists. The results of this study revealed the limited ability of this artificial model to determine the causer of an event. The model achieved an average F1-Score of 0.53 in identifying the event-relevant word and an average F1-Score of 0.63 in determining the causer of the event, compared to human performance. This performance indicates that the ChatGPT model, as one of the large language models, while demonstrating text generation capabilities approaching human-level performance, still lacks the ability to understand and comprehend semantic relationships within a sentence to the same extent as humans do. Nevertheless, the study suggests utilizing this model in semantic role labelling with human intervention by presenting the results to a specialist for review and correction, which may aid in data annotation for machine learning.

Keywords: Event, Thematic Roles, Causer of the Event, Large Language Model (LLMs), ChatGPT.

***المؤلف المراسل:** يحيى بن علي آل مريع عسيري

البريد الالكتروني الرسمي: yaasseri@kku.edu.sa

Doi: https://doi.org/10.54940/ll90229146

1658-8126 / © 2024 by the Authors.

Published by ***J. Umm Al-Qura Univ. Lang. Sci. and Lit.***

مقدمة:

منذ بداية التفكير في ذكاء الآلة كانت مسألة اللغة حاضرة في وعي علماء الذكاء الاصطناعي بوصفها إحدى القدرات الإنسانية العليا التي يقاس بها الذكاء، ففي منتصف القرن الماضي عدَّ عالم الرياضيات ومؤسس الذكاء الاصطناعي ألن تورنغ (Alan Turing) امتلاك الآلة قدرةً على الحوار مع الإنسان بلغته البشرية دليلا على امتلاك الآلة ذكاء يقارب ذكاء الإنسان أو يساويه، ومنذ ذلك الحين والعمل على اكتشاف آلية عمل اللغة البشرية وتطوير قدرة لغوية للآلة تحاكي تلك القدرة البشرية يحتل مركز اهتمام بعض علماء اللسانيات وعلماء الحاسوب، وقد نتج عن ذلك ظهور مجالات معرفية متعالقة كمعالجة اللغات الطبيعية واللسانيات الحاسوبية وتكنولوجيات اللغات الطبيعية التي تشترك جميعها في هدف واحد هو تطوير قدرة الآلة على فهم الكلام البشري وتوليد بصورة تحاكي أو تقارب قدرة الإنسان. ولئن كان التطور في مجال معالجة اللغة على المستوى النصي/الشكلي قد حقق نجاحات كبيرة فإن التقدم على مستوى إدراك الآلة للمعنى وفهمه على المستويين الدلالي والتداولي ما زال يشكل تحديًا كبيرًا لما تحتاجه معالجة المعنى من معرفة معجمية/ أنطولوجية ولما يتسم به جانب المعنى من طبيعة تجريدية وكذلك لما يصاحبه من غموض دلالي وتداولي لا سبيل إلى التغلب عليه دون الاستعانة بمعلومات ذات صلة بالسياق بمعناه النصي الضيق أو الفيزيائي الموسع.

والمنطلق النظري لهذه الدراسة التطبيقية هو ما يعرف في الدرس اللساني بنظرية الأدوار المحورية (Theta-role theory)، المعروفة أيضًا باسم نظرية الأدوار الدلالية، وهي مفهوم أساسي في الدلالة اللغوية يهدف إلى فهم العلاقات الدلالية بين الأفعال/المحمولات ومتعلقاتها/ أو ما يعرف بالحدود (Arguments)، أو بمعنى أخر إلى فهم العلاقة بين البنية الدلالية للفعل والبنية النحوية. وتوفر هذه النظرية إطارًا معرفيًا لفهم كيفية تفاعل المعجم مع النحو وكيف يسهم هذا التفاعل في فهم دلالة الجملة حيث تفترض هذه النظرية أن لكل فعل متعلقات محددة تختلف باختلاف المعاني التي يحملها، وأن لكل متعلق من هذه المتعلقات دوًا محوريا محددًا لا يمكن وصفه من خلال العلاقات النحوية (الفاعلية والمفعولية، إلخ) فقط. ويمكن توضيح مفهوم الأدوار المحورية من خلال "كسر الولدُ النافذةَ". فـ"الولد" في هذه الجملة الذي يحتل موضع الفاعل يحمل دور المنفذ (Agent) للحدث الذي يدل عليه المعنى الفيزيائي للفعل "كَسَرَ"، بينما "النافذةَ" التي تحتل موضع المفعول به تحمل دور الضحية (Patient) أو المتأثر بالحدث، ولكن ولو قلنا:" انكسرت النافذة"، فسنلاحظ أن "النافذة" تحتل موضع الفاعل مع أنها تحمل دور الضحية وليس المنفذ للحدث.وقد طُورت هذه النظرية على أيدي بعض علماء اللسانيات مثل تشارلز ج. فيلمور (Charles. J. Fillmore) وتشومسكي (Chomsky)، وداوتي (Dowty)، (1,2). واكتسبتْ شهرة في النصف الثاني من القرن العشرين، وما تزال مجالًا مهما من مجالات الدرس اللساني.

وتُعد معرفة الآلة للأدوار المحورية للمشاركين في الحدث الذي تحيل عليه الجملة، عبر الأفعال أوما يشبهها من الأسماء المشتقة، أحد الجوانب المتعلقة بفهم دلالة الجملة الذي لقي اهتمامًا كبيرًا من المتخصصين في مجال اللسانيات الحاسوبية وتكنولوجيا اللغات الطبيعية (3)، وقد نشأ هذا الاهتمام المتزايد بهذا الجانب نتيجة الإدراك بأن التحليل النحوي للجملة، كما هو في البنوك الشجرية وغيرها من نماذج التحليل النحوي، لا يحقق الحد الأدنى من مستوى فهم دلالة الجملة كإدراك المعلومات المتعلقة بالمشاركين في الحدث وأدوارهم الدلالية والإجابة على أسئلة من قبيل: من الذي قام/ تسبب في الحدث، ومن الذي وقع عليه الحدث أو تأثر به؟ وأين ومتى وكيف ولماذا وقع الحدث؟ وهي أسئلة تتعلق بجوانب مهمة من دلالة الجملة تساعد الآلة على استنتاج كثير من المعلومات الدلالية الأخرى. وإذا كان التحليل النحوي لا يجيب عن ذلك فقد كان من الضروري العمل على تطوير خوارزميات تقوم بتحديد الأدوار المحورية/الدلالية التي يُشفِّرها الفعل في بنيته المعجمية، وقد بدأ التفكير في هذه المسألة بإنشاء مصادر معجمية للأفعال تمثل موارد حاسوبية يُستعان بها على تدريب الآلة لمعرفةِ ما يتعلق بالأفعال من معلومات دلالية لا سبيل إلى معرفتها دون الاستعانة بهذه الموارد المعجمية، وقد استُعملتْ هذه المعرفة المعجمية في تدريب خوارزميات تعلم الآلة لتحديد الأدوار المحورية، وسيتضح لنا في جزء قادم من هذه الدراسة أن الجهود الحاسوبية المهتمة بهذا المجال قد قطعتْ شوطًا كبيرًا في المجالين النظري والتطبيقي مستعينة بتطور خوارزميات الذكاء الاصطناعي المشرَف عليها (Supervised Learning) التي تُدرب على بيانات محدودة لتحديد الأدوار الدلالية/المحورية، وهو ما يعرف بخوارزميات تحديد الأدوار الدلالية (Semantic Role Labeling).

ومع تطور خوارزميات الذكاء الاصطناعي في السنوات الخمس الأخيرة وظهور ما يسمى بالنماذج اللغوية الكبيرة (Large Language Models) كنموذج الـ(ChatGPT)، تَوجَّه اهتمامُ الباحثين في مجال معالجة اللغات الطبيعية إلى اختبار قدرة هذه النماذج الضخمة على القيام ببعض المهمات التي تتطلب معرفةً لغوية؛ وذلك للكشف عن إمكانية استخدام هذه النماذج بدلا من استخدام الطرق التقليدية في معالجة اللغات الطبيعية المقتصرة على بناء خوارزميات خاصة موجهة لمهام خاصة كالخوارزميات المشار إليها آنفًا. والإسهام الذي تقدمه هذه الدراسة يأتي في هذا السياق، فهي تحاول أن تكشف عن قدرة النماذج اللغوية الضخمة، وتحديدًا نموذج (ChatGPT)، على معرفة المتسبب في الحدث دون الحاجة إلى الاستعانة بخوارزميات التحليل الصرفي والنحوي وخوارزميات التحليل الدلالي ودون الاستعانة بالموارد المعجمية المطلوبة لذلك، فإذا كانت الخوارزميات التقليدية المستخدمة لتحديد الأدوار المحورية، ومنها المتسبب في الحدث، تشترط كون الجمل المدخلة محللة نحويًا وتشترط كذلك توفر موارد معجمية خاصة بالأفعال ومشتقاتها وما تشفره من معلومات معجمية، فإن استخدام النماذج اللغوية الضحمة لا يحتاج إلى ذلك كله؛ ومن هنا يأتي إسهام هذه الدراسة في هذا المجال؛ إذ لم يسبق، حسب علم الباحث، اختبار قدرة النماذج اللغوية الضحمة على معرفة المتسبب في الحدث. وقد تكونتْ بنيةُ هذه الدراسة من عدة مباحث رئيسة، بدأتْ بتفصيل القول في مشكلة الدراسة وحدودها وأهدافها وهو ما نجده في المبحث (2)، وقد اقتضت

الطبيعة العلمية لهذه الدراسة أن يعقب ذلك مباشرة ملخصٌ للدراسات السابقة بما يكشف عن المساهمة العلمية التي تقدمها هذه الدراسة في هذا السياق العلمي، وهو ما نجده في المبحث (3)، ثم جاء بعد ذلك الحديث مفصلا في المبحث (4) عن المنهجية المتبعة لاختبار قدرة نموذج الـ(ChatGPT)على معرفة المتسبب في الحدث والإجراءات التجريبية المتبعة لقياس دقته ومقارنته بالأداء الإنساني، وقد جاء المبحث (5) لمناقشة نتائج هذه الدراسة وتقديم المقترحات.

مشكلة الدراسة وحدودها وأهدافها:

مشكلة الدراسة:

من المشكلات التي تواجه نماذج الذكاء اللغوي الاصطناعي مشكلة القدرة على فهم اللغات الطبيعية، وحينما نقول فهم اللغات الطبيعية فنحن نقصد به القدرة على فهم المعنى بشقيه الدلالي والتداولي؛ فتفاعل الآلة مع الإنسان وإنجازها لكثير من المهام المطلوبة يرتكز على قدرة هذه الآلة على التواصل معه بلغته الطبيعية وقدرتها على فهم المراد والتحليل والاستنتاج، وهو هدف تسعى إليه كثير من تطبيقات معالجة اللغات الطبيعية. وما تزال مشكلة فهم المعنى الدلالي للجملة تمثل تحديًا كبيرًا لكثير من الجهود المبذولة في هذه السياق، فكثير من المعلومات المهمة التي تُشفِّرُها الجملة لا يمكن استنباطها من البنية السطحية/ الظاهرة للجملة، ومن ذلك ما يتعلق بمعرفة الأدوار المحورية للمشاركين في الحدث الذي تشفره الأفعال وما يشبهها من المشتقات الدالة على الحدث. ونقصد بالمشاركين في الحدث: من قام/ تسبب في الحدث ومن وتأثر بالحدث والمكان والزمان. والمشكلة التي تقاربها هذه الدراسة، وأعني بها هنا معرفة المتسبب في الحدث كأحد الأدوار المحورية، تأتي في هذا السياق، ولتوضيح ذلك نعطي الأمثلة الآتية:

1- كسر الولد زجاج السيارة.

2- تحطم زجاج السيارة من شدة الحرارة.

3- قُتِل الكثيرُ على أيدي المستعمرين.

لنفترض أن هذه الجمل أُعطيت للإنسان وطلب منه أن يجيب على سؤال واحد وهو من الذي قام أو تسبب في الحدث الذي تدل عليه كل جملة من الجمل السابقة «حدث الكسر»، و«حدث التحطيم»، و«حدث القتل»، سوف نلاحظ أنه من السهل بالنسبة للإنسان أن يجيب عن هذا السؤال وأن يقوم بتحديد المسؤول أو المتسبب في الحدث في كل جملة من هذه الجمل، وإجابته تلك تخفي وراءها معرفة لغوية تتمثل في: 1) قدرته على معرفة الكلمة التي تدل على الحدث في كل جملة، و2) قدرته على معرفة المعنى الخاص بالفعل في سياق الجملة وتمييزه عن المعاني الأخرى التي للفعل الموجودة في الذاكرة المعجمية، فالفعل "كَسَرَ"، على سبيل المثال، له معان معجمية متعددة أحدها هو الدلالة على تغير الحالة الفيزيائية لشيء ما، ولكل معنى بنيته المحورية، و3) قدرته على معرفة المتسبب في هذا الحدث (المنفذ أو السبب أو المثير) المرتبط بهذا المعنى المشفر في معجمه الذهني، 4) قدرته على تمييز المركبات النحوية المكونة للجملة التي يرد فيها الفعل (الفاعل والمفعول به والجار والمجرور، وغيرها)، و5) قدرته على الربط بين الدور المحوري/الدلالي المشفر في بنية الفعل المعجمية وبين المكون النحوي الذي يعبر عنه في الجملة، وهو ما نعبر عنه في هذه الدراسة بمعرفة المتسبب في الحدث. ولا يخفى على المتأمل لهذه الأمثلة أن المتسبب في الحدث قد يرتبط بالفاعل كما في الجملة (1)، وقد يرتبط بالاسم المجرور كما في (2)، وقد يعبر عنه بالمضاف إليه كما في (3). والجدول رقم (1) يبين المكونات النحوية لكل جملة وما يرتبط بها من أدوار محورية.

الجدول (1): أمثلة للتفريق بين التحليل النحوي والدلالي لمتعلقات الفعل.

الجملة	معنى الفعل	المكون النحوي	الدور المحوري
1- كسر الولد زجاج السيارة	1	الفاعل = الولد	المنفذ
		المفعول به = زجاج السيارة	الضحية
2- تحطم زجاج السيارة من شدة الحرارة	1	الفاعل = زجاج السيارة	الضحية
		اسم مجرور = شدة الحرارة	السبب
3- قُتل الكثيرُ على أيدي المستعمرين	1	نائب الفاعل = الكثير	الضحية
		اسم مجرور = أيدي المستعمرين	المنفذ

ومع كل هذا، فإن هذه المسألة، وأعني بها مسألة الإجابة عن سؤال من المتسبب في الحدث، قد تبدو سهلة ومباشرة بالنسبة للإنسان، وقد لا يستغرق ذلك منه وقتًا كبيرًا في تحديد المسؤول عن الحدث في أي جملة من هذه الجمل؛ وذلك لاشتمال معرفته اللغوية التي تمثل نظامه اللغوي على المعرفة الدلالية المشار إليها، لكن الأمر بالنسبة للآلة يبدو غير ذلك، فمعرفة المتسبب في الحدث في الجمل السابقة تحتاج إلى: 1) موارد معجمية للأفعال تحتوي على رصد للأفعال وما يشبهها من المشتقات الدالة على الحدث ورصد للمعاني التي يدل عليها كل فعل مع تحديد الأدوار المحورية المرتبطة ببنية الموضوعات التي يتطلبها الفعل لكل معنى من معانيه، ومن الأمثلة على تلك الموارد المعجمية ما يُعرف بشبكة الأفعال وبنك المحمولات وشبكة الأطر، وهي موارد حاسوبية طورت لغرض تزويد الآلة بهذه المعلومات المعجمية المتعلقة بمعاني الفعل وأدواره المحورية، 2) خوارزميات تقوم بالتحليل الصرفي والتحليل النحوي لتحديد المكونات النحوية، و3) خوارزميات لتحديد الأدوار المحورية (Semantic role labelling) تقوم باختيار معنى الفعل المناسب للسياق بالاستعانة بالموارد المعجمية وربط الأدوار المحورية المحددة مسبقًا في تلك الموارد بما يقابلها من المكونات النحوية، وهذه الخطوات المتسلسلة لتعرف الآلة على الأدوار المحورية تمثل الطريقة الشائعة والمتعارف عليها بين المتخصصين.

والسؤال الذي تحاول أن تجيب عنه هذه الدراسة هو سؤال يأتي في سياق سؤال أعم منه، وهو هل تستطيع النماذج اللغوية الضخمة مثل نموذج (ChatGPT) تحديد الأدوار المحورية دون الحاجة إلى الاستعانة بخوارزميات التحليل الصرفي والنحوي وخوارزميات التحليل الدلالي ودون الاستعانة بالموارد المعجمية المطلوبة لذلك؟ وفي سياق أضيق يتناسب مع ضيق المساحة التي تتيحها لنا هذه الدراسة، فإن السؤال هو: هل يستطيع هذا

النموذج الاصطناعي تحديد المسؤول عن الحدث أو المتسبب فيه بمجرد إعطائه مجموعة من الجمل وسؤاله بالطريقة التي يسأل بها الإنسان دون الاستعانة بتلك الموارد الحاسوبية وتلك الخوارزميات؟ وإذا كان بمقدوره ذلك، فإلى أي درجة تصل دقة هذا النموذج في الإجابة على هذا السؤال مقارنة بالإنسان؟ وكيف يمكننا قياس ذلك؟

حدود الدراسة ومجالها:

تندرج هذه الدراسة ضمن مجال الدراسات اللسانية الحاسوبية المهتمة باختبار قدرة الآلة على فهم اللغات الطبيعية، وتحديدًا قدرة الآلة على معرفة الأدوار المحورية للمشاركين في الحدث الذي يشفره الفعل أو ما يشبهه من الأسماء المشتقة. وفيما يأتي تحديد للمصطلحات التي تدور حولها هذه الدراسة:

1- المتسبب في الحدث (Causer of the Event): يقصد بالمتسبب في الحدث هنا كل ما من شأنه القيام بدور المنفذ للحدث (Agent) أو السبب (Cause) أو المثير (Stimulus)، أو الأداة (Instrument)، وهي مجموعة الأدوار المحورية المعروفة بـ(Proto-Agent Roles) كما سماها عالم اللسانيات داوتي (4)، وهي ما يُعبر عنه في بنك المحمولات (5) بالرمز (ARG0)، ويجمعها كونها أدوار مسؤولة عن إحداث التغير الفيزيائي أو النفسي أو المكاني للضحية أو المستفيد أو الموضوع، كما توضحه الأمثلة الآتية:

- كسر الولد [منفذ] النافذة [ضحية].
- فزع الولد [مجرب] من الأسد [سبب].
- أنقذ الحبل [أداة] الغريق [مستفيد].

الجدول (2): مقولة المتسبب في الحدث، وما يقابلها في بنك المحمولات وشبكة الأفعال.

مقولة المتسبب في الحدث وما يقابلها في النماذج المختلفة			
في هذه الدراسة	النماذج اللسانية	بنك المحمولات	شبكة الأفعال
المتسبب في الحدث	Proto -Agent roles (الأدوار المشابهة للمنفذ)	ARG0	المنفذ (Agent) – له إرادة وقدرة
			المثير (Stimulus) – له أثر نفسي/ذهني
			السبب (Cause) – من عوامل الطبيعة
			الأداة (Instrument) – وسيلة لتنفيذ الحدث

2- الحدث (Event): يقصد به في هذه الدراسة الحدث الذي يدل عليه الفعل أو ما يقوم مقامه من الأسماء المشتقة الدالة على الحدث، كحدث «الكسر»، أو «الفزع»، أو «الإنقاذ» في هذه الأمثلة.

3- النماذج اللغوية الضخمة (LLMs): هي أنظمة من أنظمة الذكاء الاصطناعي المتطورة صُممت باستخدام خوارزميات الشبكات العصبية وتحديدًا نماذج التعلم العميق، لغرض فهم وتوليد اللغة البشرية، وقد تم ذلك بتدريبها على كميات ضخمة من نصوص تتألف من مليارات أو حتى تريليونات الكلمات بحيث يمكنها حساب تتابع/ توالي الكلمات في الجمل المستعملة في لغة ما، وهو ما نتج عنه القدرة على التفاعل مع الإنسان بالرد على طلباته واستفساراته بتوليد النصوص والقيام بكثير من المهام اللغوية كالبحث والتلخيص والإجابة على الأسئلة والترجمة الآلية وغيرها من المهام(6).

4- نموذج (ChatGPT): وهو أحد تطبيقات النماذج اللغوية الضخمة (LLMs) التابع لشركة (OpenAI) المهتمة بالذكاء الاصطناعي(10–7)، وقد صُمّم لغرض الحوار والتفاعل مع الإنسان بلغته الطبيعية، وهو نموذج يعتمد على ما يعرف بـ(Generative pre-trained Transformer) وتعني المحولات التوليدية المدربة مسبقًا على بيانات(*)، وقد استخدمت النسخة (ChatGPT-3.5) في هذه الدراسة.

أهداف الدراسة:

من خلال العرض السابق لمشكلة الدراسة، يمكن تلخيص أهدافها في فيما يأتي:

1- اختبار قدرة نموذج الشات جي بي تي (ChatGPT) على تحديد الكلمة الدالة على الحدث في الجملة، وهذا متطلب أساسي لمعرفة المتسبب في الحدث، فلكي نختبر قدرته على معرفة المتسبب في الحدث، لابد من معرفة الكلمة الدالة على الحدث أولا.

2- اختبار قدرة نموذج الشات جي بي تي (ChatGPT) على معرفة/ تحديد المتسبب في الحدث في الجملة.

3- تقييم دقة هذا النموذج بمقارنته بأداء الإنسان الذي أعطيت له الجمل نفسها وطلب منه تحديد الكلمة الدالة على الحدث وتحديد المتسبب في الحدث.

وقبل البدء في مقاربة هذه الأهداف، سيتولى القسم الآتي من هذه الدراسة تقديم ملخص للدراسات السابقة في هذا المجال بما يسمح بمَوْضَعَة هذه الدراسة في سياقها العلمي وبيان ما يمكن أن تضيفه في هذا المسار البحثي.

الدراسات السابقة:

احتل تحديد الأدوار المحورية للمشاركين في الحدث (Semantic Role Labelling) مكانة كبيرة من اهتمام المتخصصين في مجال اللسانيات الحاسوبية وتكنولوجيا اللغات الطبيعية وتحديدًا عند أولئك المهتمين بالتحليل الدلالي على مستوى الجملة (3)، وقد نشأ هذا الاهتمام المتزايد بهذا الجانب نتيجة الإدراك بأن التحليل النحوي للجملة، كما هو في البنوك الشجرية وغيرها من نماذج التحليل النحوي، لا يحقق الحد الأدنى من مستوى فهم دلالة الجملة كإدراك المعلومات المتعلقة بالمشاركين في الحدث وأدوارهم الدلالية والإجابة على أسئلة من قبيل: من الذي قام/ تسبب في الحدث، ومن الذي وقع عليه الحدث أو تأثر به؟ وأين ومتى وكيف ولماذا وقع الحدث؟ وهي أسئلة تتعلق بجوانب مهمة من دلالة الجملة وتساعد الآلة على استنتاج كثير من المعلومات الدلالية الأخرى، وهو ما

(*) لمزيد من الاطلاع يمكن الرجوع إلى: https://chat.openai.com/auth/login

يعد أساسًا ومنطلقًا لكثير من التطبيقات الحاسوبية، مثل الإجابة على الأسئلة (11,12) أو استخلاص المعلومات من النصوص(13,14) والتلخيص(15) وقياس التشابهات الدلالية بين الجمل (16) واستخلاص شبكة المعرفة الدلالية (17) وبناء الأنطولوجيا المعجمية (18) وإزالة الغموض المعجمي (19) وتعليم الروبوتات (20)، وغيرها من التطبيقات، فالآلة لا يمكنها استنباط هذه المعلومات من الجملة إذا لم تستطع تحديد الكلمات الدالة على الأحداث وتحديد المشاركين فيها ومعرفة الدور الذي يقوم به كل مشارك في هذا الحدث.

وإذا كان التحليل النحوي، كما أشرنا، لا يجيب عن ذلك فقد كان من الضروري العمل على تطوير خوارزميات تقوم بتحديد الأدوار المحورية/الدلالية التي يشفرها الفعل في بنيته المعجمية. وقد بدأ التفكير في هذه المسألة بإنشاء مصادر معجمية للأفعال تمثل موارد حاسوبية لهذه التطبيقات بحيث يستعان بها في معرفة ما يتعلق بالأفعال من معلومات دلالية لا سبيل إلى معرفتها دون الاستعانة بهذه المصادر المعجمية، ومن هذه الموارد الحاسوبية التي أُعدت لهذا الغرض ما يعرف بشبكة الأطر (21–27) وشبكة الأفعال (28–30) وبنك المحمولات (5,31–33). والهدف من هذه الموارد الحاسوبية يتمثل في 1) رصد الأفعال وتمييز معاني الفعل الواحد بعضها عن بعض و2) تحديد الأدوار الدلالية المرتبطة بهذا المعنى أو ذاك من معاني الفعل وهو ما يسمى بتحديد بنية الموضوعات أو البنية المحورية (argument structure)، وهي ما يمكن أن نسميها بمتعلقات الفعل، فلكل فعل بنيته المحورية التي يتطلبها المعنى الذي يدل عليه. ولئن اختلفت هذه الموارد الحاسوبية في أهدافها وطرائق بنائها وتمثيلاتها إلا أنها تشترك جميعًا في كونها موجهة لغرض تعريف الآلة بالأدوار المحورية المشفرة في البنية المعجمية لكل فعل، وبمعنى أدق الأدوار الدلالية المشفرة في البنية المعجمية لكل معنى من معاني الفعل. ولأن اللغات لها خصوصياتها المعجمية، فقد بني بنك للمحمولات العربية (32,33) وشبكة أفعال للعربية (34) وكذلك شبكة أطر الأفعال العربية (23). وقد تفاوتت هذه الموارد المعجمية في تصنيفها الأدوارَ الدلالية، فمنها ما اكتفى بتوسيمات (labels) عامة مثل بنك المحمولات الذي استخدم الواسمات (ARG6-ARG0)، ومنها ما كان أكثر تفصيلا وأكثر دقة مثل شبكة الأفعال حيث استخدمت ما يقرب من 24 دورًا دلاليًا منها المنفذ (Agent) والسبب (Cause) والمُحفز/ المثير (Stimulus) والأداء (Instrument) والضحية (Patient) والمُعاني/ المجرب (Experiencer) والمستفيد (Beneficiary) والمستقبل (Recipient) والوجهة (Destination) والمصدر (Source) والمكان (Location)، وغيرها.

وقد استخدمت هذه الموارد الحاسوبية للتحشية اليدوية (Manual annotation) لعدد كبير من الجمل المحللة نحويا، وذلك بتوسيم/تحديد الدور المحوري لكل مكون نحوي يمثل أحد متعلقات الفعل الدال على الحدث أو الاسم المشتق منه. وفي عملية التوسيم يقوم المُوسِّم بالاستعانة بأحد تلك الموارد المعجمية لاختيار المعنى المناسب للجملة من بين معاني الفعل الموجود في الجملة، ثم توسيم الأدوار المحورية لكل مكون نحوي من مكونات الجملة بالاستعانة بالإطار أو الأدوار المحورية المحددة مسبقا في أحد هذه الموارد المشار إليها سابقًا، وبعد الانتهاء من توسيم عدد كبير من الجمل، يتم استخدام خوارزميات تعلم الآلة في التدريب ثم اختبارها على بيانات جديدة لم يتم توسيمها من قبل، ومن ثَمَّ تُقاس قدرةُ هذه الآلة على تحديد دور كل مشارك ومقارنة أدائها مع نموذج معياري تم توسيمه يدويا. وتُصنف خوارزميات تحديد الأدوار المحورية إلى ثلاثة أنواع: خوارزميات تعتمد على الطريقة القديمة لتعلم الآلة وهي الطريقة المعتمدة على السمات (35–40) والنوع الثاني من هذه الخوارزميات هي تلك المعتمدة على الشبكات العصبية (41–43)، وأما النوع الثالث فهو النوع الذي يمزج بين الطريقتين السابقتين (44). وقد كان للغة العربية نصيب من تلك المحاولات الجادة؛ فقد أنشئت الموارد الحاسوبية اللازمة، كما أشرنا، وطورت خوارزميات خاصة بتوسيم الأدوار الدلالية لجمل ونصوص من اللغة العربية الفصحى(37,39,45,46).

ومع تطور خوارزميات الذكاء الاصطناعي وظهور النماذج اللغوية الكبيرة كنموذج الـ(ChatGPT)، تَوجَّه اهتمامُ الباحثين في مجال معالجة اللغات الطبيعية إلى الاستفادة من هذه النماذج الضخمة في توسيم البيانات اللغوية على جميع المستويات، ومنها المستوى الدلالي، بدلا من استخدام الطرق التقليدية، وأعني طريقة بناء خوارزميات خاصة موجهة لمهام خاصة كالخوارزميات المشار إليها آنفًا. والمساهمة التي تقدمها هذه الدراسة تأتي في هذا السياق، فهي تحاول أن تكشف عن قدرة النماذج اللغوية الكبيرة، وتحديدًا نموذج (ChatGPT)، على معرفة المتسبب في الحدث دون الحاجة إلى الاستعانة بخوارزميات التحليل الصرفي والنحوي وخوارزميات التحليل الدلالي ودون الاستعانة بالموارد المعجمية المطلوبة لذلك. والسؤال هو: هل يستطيع هذا النموذج تحديد المسؤول عن الحدث أو المتسبب فيه بمجرد إعطائه مجموعة من الجمل وسؤاله بالطريقة التي يسأل بها الإنسان دون الاستعانة بتلك الموارد الحاسوبية وتلك الخوارزميات؟ وإلى أي درجة تصل دقة هذا النموذج في الإجابة على هذا السؤال مقارنة بالإنسان؟ وبهذا تأتي المساهمة العلمية لهذه الدراسة إذ لم يسبق –حسب علم الباحث– أن وجدت دراسة تجيب عن هذه الأسئلة.

طريقة اختبار قدرة (ChatGPT) على معرفة المتسبب في الحدث.

أشرنا في موضع سابق من هذه الدراسة إلى أن نماذج الذكاء الاصطناعي اللغوية الضخمة، كنموذج الـ(ChatGPT)، أصبحت تستخدم من قبل المتخصصين في مجالي معالجة اللغات الطبيعية واللسانيات الحاسوبية لأغراض بحثية وتطبيقية على حد سواء، ومن ذلك استخدامها في عملية التوسيم والتصنيف الآلي بديلا عن الخوارزميات التقليدية المصممة لمهام محددة والمدربة على بيانات قليلة، وهدف هذه الدراسة كما أسلفنا يندرج في هذا السياق. ويأتي هذا المبحث لشرح المنهجية المتَّبعة في هذه الدراسة لاختبار قدرة نموذج الـ(ChatGPT) على معرفة المتسبب في الحدث. وملخص هذا المبحث هو أننا سنقوم باختباره على عدد 200 جملة من

الجمل العربية الفصيحة التي تحتوي على أفعال أو مشتقات تتطلب في بنيتها المعجمية وجود متسبب في الحدث، ثم مقارنة أدائه بأداء الإنسان، وتتكون هذه التجربة من أربع خطوات: 1) جمع وإعداد البيانات اللغوية وتهيئتها، و2) توسيم هذه البيانات يدويًّا، و3) هندسة الأوامر/المدخلات (prompt engineering) المناسبة لطبيعة الموضوع، و4) تقييم النتائج بمقارنة أداء هذا النموذج بالبيانات الموسّمة يدويًّا، وفي الفقرات الآتية شرح مفصل لهذه الخطوات:

إعداد البيانات/المادة اللغوية وتهيئتها:

أول خطوة في بناء هذه التجربة هي جمع الجمل التي تحتوي على حدث له متسبب، ولكن لا بد أن يسبق عمليةَ جمع الجمل وتحديدها اختيارُ الأفعال التي تدل على هذه الأحداث التي لها متسبب، فمن المعلوم أن الأفعال التي ينطبق عليها هذا الشرط هي مجموعة جزئية من أفعال العربية، أي فقط تلك الأفعال التي تتطلب في بنيتها المعجمية وجود منفذ أو مسبب أو مثير للحدث، وليس أي فعل من أفعال العربية يدل على ذلك. ولو أخذنا مثالا على ذلك الفعل (كَسَرَ) بمعنى أحدث تغييرًا في الحالة الفيزيائية لشيء محسوس فإن دلالته المشفرة في الذهن تتكون من (متسبب) في الكسر و(الشيء المكسور أو الضحية)، لكن لو أخذنا الفعل (سَمِعَ) فإن هذا الفعل لا يدل على وجود منفذ متسبب في الحدث، ومثله (رَكَضَ)، ولهذا سيكون تركيزنا في هذه الدراسة على الأفعال التي تحتوي في بنيتها الدلالية التصورية على متسبب في الحدث، وفق التعريف المذكور سابقا، يتعدى أثره إلى مشارك آخر فيغير من حالته. وهذه الأفعال هي من طبقات الأفعال الدالة على تغيير الحالة الفيزيائية أو تغيير الحالة النفسية أو الحالة المكانية. وقد اتضح في موضع سابق أن الأدوار المحورية لكل معنى من معاني الفعل هي من قبيل المعلومات الدلالية المشفرة في البنية المعجمية الذهنية للإنسان؛ ولذلك نستعين بشبكة الأفعال العربية لاستخلاص هذه النوع من الأفعال والأطر الدلالية التي تمثل أحد صيغ بنية الموضوعات الموضحة في الجدول (3).

بنية الموضوعات/ الأدوار المحورية	أمثلة
[منفذ – ضحية] [Agent – Patient]	كَسَرَ، أَحْرَقَ، هَدَمَ، قَتَلَ، قَطَعَ، طَعَن، اغتالَ
[محفز – مجرب] [Stimulus – Experiencer]	أثار، أفزَعَ، خوّف، أبْكى، أفْهم، هَدّد،
[منفذ – موضوع -- مكان] [Agent – Theme-- Location]	أبْعـدَ، أخْـرجَ، أغْلـقَ، طَـرَدَ، هَـرَّبَ، قَصَفَ، أطلَقَ
[سبب – موضوع] [Cause-Theme]	عصَفَ، زاد، خَفّضَ، عَلَّق

وبعد تحديد الأفعال وفقا لهذه الشروط الموضحة، استعانتْ هذه الدراسة بمدونة العربية المعاصرة التابعة لمدينة الملك عبد العزيز للعلوم والتقنية(*)، وبعض المقالات الصحفية لجمع الجمل المطلوبة، ولكون الفعل الواحد قد يُستعملُ في أكثر من تركيب نحوي، وهو ما قد يختلف معه موضع المتسبب في الحدث، سمحنا بتكرار الأفعال دون تكرار التراكيب، وقد تعاملنا مع مشتقات تلك الأفعال في حال غياب الصيغة الفعلية كبديل عن الفعل لكونها، أعني المشتقات، تحتفظ ببنية الموضوعات وبالدلالة على الحدث الذي يدل عليه الفعل، وقد نتج عن ذلك مدونة مكونة من 200 جملة وتشتمل على 43 فعلا/مشتقًا. ولكون هذه الدراسة ليست إلا دراسة أولية يقصد منها أن تكون منطلقا لغيرها، تجنبنا الجمل المعقدة.

توسيم البيانات/ الجمل يدويًّا:

لقياس قدرة الـ(ChatGPT) على معرفة المتسبب في الحدث نحتاج إلى مقارنة أداء هذا النموذج ببيانات مُوسَّمة يدويًّا، ويُقصد بالتوسيم هنا، حسب المتعارف عليه في المدونات الحاسوبية، إضافة معلوماتٍ لغوية على النص/الجملة لمساعدة الآلة على إدراك المعلومات اللغوية أو المعاني اللغوية التي لا تظهر في البنية السطحية. وبعبارة أخرى، يمكن القول: إن التوسيم هو عملية الربط بين مكون لفظي وبين مدلول أو معنى ينتمي إلى مقولة أو صنف دلالي أو قواعدي. فإذا كانت الأدوار الدلالية/ الأدوار المحورية هي مقولات أو أصناف دلالية، فعملية توسيم الأدوار الدلالية هي عبارة عن تحديد المكونات النحوية المتعلقة ببنية الفعل المعجمية وربط كل منها بالصنف أو الدور الدلالي المناسب له، وهو صنف أو مقولة دلالية محددة مسبقا.

الجدول (3): بنية موضوعات الأفعال الدالة على متسبب في الحدث.

الجدول (4): جدول يوضح هندسة الأوامر (prompt engineering) المستخدمة في هذه التجربة.

اسم الصيغة	الصيغة	الهدف منها
---	اتبع التعليمات الآتية:	– إعطاء إرشادات عامة.
Persona Pattern	1- تصرف وكأنك لغوي.	– تضييق مجال البحث.
Meta Language Creation Pattern	2- حدد الكلمة الدالة على الحدث في الجمل العربية المعطاة.	– تحديد المطلوب بدقة.
Meta Language Creation Pattern	3- حدد المتسبب في الحدث في كل جملة من هذه الجمل.	– تحديد المطلوب بدقة.
Template Pattern	4- أعطني الإجابة على الصيغة الآتية مستعملا الفواصل للفصل بينها: «الجملة»، «الكلمة الدالة على الحدث»، «المتسبب في الحدث».	– تحديد الصيغة التي تكون عليها المخرجات وترتيب المعلومات المطلوبة.
Visualization Generator Pattern	5- أحفظ الإجابات على شكل ملف (CSV).	– طريقة حفظ المخرجات لتسهيل التعامل معها حاسوبيا.

(*) لمزيد من الاطلاع يمكن الرجوع إلى: https://corpus.kacst.edu.sa/more_info.jsp

ومع أن لكل فعل بنيته الموضوعاتية المعجمية التي قد تشتمل على أكثر من دور محوري، وقد يصل بعضها إلى أربعة أدوار محورية، إلا أن التوسيم الذي سنقوم به في هذه الدراسة سوف يقتصر على أمرين: 1) تحديد الكلمة الدالة على الحدث و2) تحديد المتسبب في الحدث. وعلى هذا الأساس، فمهمة التوسيم اليدوي هنا هي عملية الربط بين مكون نحوي من مكونات الجملة وبين إحدى المقولات الدلالية التي تنتمي إلى الأدوار الدلالية، وهي مقولة المتسبب في الحدث كما عرفناها سابقا. وقد أُعْطِيت الجمل لشخصين وطلب منهما توسيمها يدويا وفقا للتعليمات ولإجراءات التوسيم الموضحة هنا، وتم مطابقة التوسيم بين هذين الشخصين، وأظهرت هذه المطابقة التوافق بينهما في كل مهمة بنسبة 100٪، وهي بلا شك مهمة يسيرة جدًا، فمن السهل بالنسبة للإنسان المتحدث بالعربية أن يحدد الكلمة/الفعل الدال على الحدث والمتسبب في الحدث، ولكنه أمر نحتاجه لمقارنة معرفة الآلة في هذا الجانب بمعرفة الإنسان.

هندسة مدخلات الـ (Prompt Engineering):

من المتعارف عليه بين المتخصصين في مجال الحوسبة والذكاء اللغوي الاصطناعي أن النماذج اللغوية الضخمة (LLMs) ذات قدرة عالية على التفاعل مع المستخدم وتوليد الجمل بصورة سليمة نحويًا ومتسقة دلاليا، سواء فيما تولده من جمل ردًا على استفسار أو بحث عن معلومة أو تلخيص لنص معين أو حتى إجراء تجربة معينة في مجال معالجة اللغات الطبيعية. ومع ما تظهره هذه النماذج من قدرة فائقة في ذلك إلا أن إحدى المشكلات الملازمة لها هي أن تحقيق المطلوب بدقة عالية يعتمد على دقة المدخلات التي يزودها به المستخدم، فكلما كانت المدخلات محددة، ساعدت هذه المدخلات تلك النماذج على فهم المطلوب بدقة ومن ثم تحقيق مخرجات ذات دقة وجودة عاليتين، ومنذ ظهور هذه النماذج بدأ المتخصصون في التفكير في كيفية التحكم بالمدخلات ووضع آليات وطرائق تساعد المستخدم على تحديد المطلوب بدقة، وقد نتج عن هذا التفكير ما يعرف بهندسة الأوامر/المدخلات (Prompt engineering)، وهي مجموعة من التعليمات والقيود العامة التي تساعد تلك النماذج اللغوية على التفاعل مع الإنسان المستخدم في مجال دلالي ضيق هو المجال الدلالي الذي يندرج تحته سؤال المستخدم أو طلبه، وقد نُشر دليل يوضح أشهر تلك الصيغ وأهمها (47). وفي هذه التجربة، اكتفينا ببعض تلك الصيغ مما يناسب المهمة المطلوبة من نموذج الشات جي بي تي (ChatGPT)، والجدول (4) يشتمل على مسرد لتلك الأوامر بصيغتها الإنجليزية وما يقابل ذلك بالعربية مع توضيح الهدف من كل أمر/ توجيه من هذه الأوامر، وهي في مجملها أوامر أو إرشادات تقوم بتنفيذ المطلوب في سياق محدد يناسب طبيعة المهمة التي تطلب منه، وقد رُتبت هذه التعليمات على النحو الموضح في الجدول، وأعطيت هذه التعليمات لهذا النموذج باللغة العربية.

جدول (5): نتائج مقارنة أداء نموذج الـ(ChatGPT) لثلاث محاولات مع حساب متوسط الدقة.

	الكلمة الدالة على الحدث			المتسبب في الحدث		
	Precision	Recall	F1-score	Precision	Recall	F-score
Exp#1	0.37	0.36	0.36	0.47	0.48	0.46
Exp#2	0.78	0.76	0.45	0.50	0.53	0.51
Exp#3	0.40	0.39	0.77	0.90	0.93	0.92
Average	0.52	0.50	**0.53**	0.62	0.65	**0.63**

مقارنة أداء النموذج مع التوسيم اليدوي:

لقياس درجة دقة أداء نموذج الذكاء الاصطناعي (ChatGPT) في معرفة المتسبب في الحدث، قارنا أداءه بنموذج يدوي معياري هو النموذج الذي سبق الحديث عنه والذي تم توسيمه والإجابة عليه من قبل الإنسان. وقد استخدمنا مقياس (F-score) لحساب دقته، وهو أحد المقاييس المعتمدة للتقييم وقياس دقة الأداء في معالجة اللغات الطبيعية. وقد عقدنا مقارنة بين نتائج ثلاث تجارب/ محاولات طُلب فيها من الـ(ChatGPT) القيام بالمهمة ذاتها وتكرارها ثلاث مرات على البيانات نفسها، وفي كل مرة يُقارن أداؤه بالنموذج الإنساني. والقصد من عقد هذه المقارنة بين أداء نموذج الذكاء الاصطناعي في المحاولات الثلاث هو حساب متوسط الدقة حيث سبقت الإشارة إلى أن إحدى مشاكل هذا النموذج عدم تطابق المخرجات (Inconsistency) عند إعادة التجربة/ إعادة التوليد. والجدول (5) يقدم ملخصًا لنتائج المحاولات الثلاث ويكشف عن مدى قدرة هذا النموذج الاصطناعي على معرفة المتسبب في ذلك الحدث بعد معرفة الكلمة الدالة على الحدث في كل جملة من الجمل المعطاة. ويتبين لنا من هذا الجدول أن متوسط أداء هذا النموذج في مهمة تحديد الكلمة الدالة على الحدث هي (F1-Score=0.53)، بينما حقق متوسط (F1-Score=0.63) في معرفة المتسبب في الحدث مقارنة بالأداء الإنساني.

جدول (6): أمثلة من أداء الـ(ChatGPT) في مهمة تحديد الكلمة الدالة على الحدث والمتسبب في الحدث.

م	الجملة	الكلمة الدالة على الحدث	المتسبب في الحدث
1	علقت وزارة الداخلية المصرية التسجيل بكلية الشرطة لمدة أسبوع	علقت	وزارة الداخلية المصرية
2	فقد تعرض للطعن في ساقه على يد أسرته	تعرض	أسرة زوجته
3	بأن مسلحين اغتالوا ضابطا كبيرا في الجيش اليمني	اغتالوا	مسلحين
4	هجوم آخر بطيارة أمريكية بدون طيار	هجوم	الطائرات الأمريكية بدون طيار
5	مقتل شقيق الرئيس الأفغاني أحمد والي على يد أحد عناصر حمايته	مقتل	أحد عناصر حمايته
6	ثارت مخاوف دول البلطيق	ثارت	غير محدد
7	وهو الأمريكي تيري جونز الذي أحرق نسخة من المصحف الشريف	أحرق	تيري جونز
8	لقي مسلحون من جنسيات أجنبية مصرعهم على يد قوات الجيش	مصرعهم	قوات الجيش
9	إطلاق النار من عناصر إرهابية	إطلاق نار	عناصر الإرهابية

مناقشة النتائج:

في المبحث السابق من هذه الدراسة قدمنا شرحًا مفصلا لتجربة قمنا فيها باختبار قدرة نموذج الـ(ChatGPT) كأحد النماذج اللغوية للذكاء الاصطناعي على معرفة المتسبب في الحدث الذي تنطوي عليه الجملة العربية المشتملة على فعل أو اسم مشتق يُشَفِّر المتكلم من خلاله هذا الحدث أو ذاك، وقد تبين لنا من خلال هذه التجربة أن هذا النموذج حقق أداء يصل متوسطه إلى (0.63) مقارنة بالمعرفة البشرية. وفي هذا المبحث من هذه الدراسة، نذكر أهم النتائج العامة التي كشفت عنها هذه التجربة:

– على الرغم مما تعكسه هذه الدراسة من أداء متدنٍ لقدرة هذا النموذج الاصطناعي على تحديد الكلمة الدالة على الحدث، إلا أنها تكشف عن قدرته على معرفة الكلمة الدالة على الحدث سواء أكانت فعلاً أم اسمًا مشتقًا، فعلى سبيل المثال، يتضح لنا من الجمل الواردة في الجدول [6] قدرته على ذلك، فالكلمات الدالة على الأحداث الواردة في الجمل (1)، (3)، (7)، (6) أفعال، بينما الكلمات الدالة على الأحداث (5)، (8)، (9) أسماء مشتقة دالة على الحدث.

– أخفق هذا النموذج في معرفة الكلمة الرئيسة الدالة على الحدث، في بعض الأمثلة، عند وجود كلمة أخرى تدل على حدث، ولكنه ليس الحدث الرئيس في الجملة، والجملة رقم (2) من الجدول نفسه تعد مثالا على هذا الإخفاق حيث عد الفعل (تَعَرَّضَ) الكلمة الدالة على الحدث الرئيس بينما الصواب هو كلمة (الطَّعْن).

– ويكشف تحليل نتائج هذه التجربة قدرة هذا النموذج الاصطناعي على تحديد المتسبب في الحدث سواء أكان اسمًا مفردًا، كما في الجملة (3) في الجدول المشار إليه أم مركبًا كما في الأمثلة (2)، (5)، (8)، (9).

– وتظهر النتائج كذلك قدرته على تحديد المتسبب في الحدث سواء أكان في موضع الفاعل (1) أم في موضع الاسم المجرور (9) أم في موضع المضاف إليه الملازم للتعبير "**على يد**" كما في (2)، (5)، (8)، وتظهر كذلك قدرته على تحديد المتسبب في الحدث الذي تقدم ذكره على الفعل كما في (3) و(7). ومن اللافت للنظر قدرته على التمييز بين عبارة "*على يد*" والمتسبب الحقيقي في الحدث وهو المضاف إليه.

– ومما كشفته هذه الدراسة قدرة هذا النموذج الاصطناعي على التعامل مع الجمل التي ذكر فيها المتسبب في الحدث والجمل أو التعابير التي لم يذكر فيها حيث أجاب عن المتسبب في الحدث للجمل التي لم يذكر فيها باستخدام عبارة "غير محدد"، متبعًا التعليمات التي أعطيت له في البداية، وهو ما نجد مثاله في الجملة (6).

– وقد تبين لنا كذلك قدرته على تحديد المتسبب في الحدث في الجمل القصيرة والجمل الطويلة كما في الجملة (5)، (8) على حد سواء دون أن تتأثر دقته بالفاصل الطويل بين الفعل والمتسبب فيه.

– وبوجه عام، تكشف هذه الدراسة أن نموذج الـ(ChatGPT)، بوصفه أحد نماذج الذكاء الاصطناعي الضخمة وإن أظهر قدرة على توليد النصوص تقارب قدرة الإنسان ما يزال غير قادر على فهم وإدراك العلاقات الدلالية داخل الجملة بصورة تقارب قدرة الإنسان على الفهم والاستنتاج.

– وأخيرًا، تقترح هذه الدراسة الاستفادة من هذا النموذج في تحليل الأدوار المحورية على أن يكون ذلك بتدخل الإنسان، وذلك بعرض النتائج على متخصص يقوم بمراجعتها وتصحيحها، وهو ما قد يساعد في توسيم البيانات تمهيدًا لاستخدامها في مهام تدريب الآلة.

الخاتمة:

ناقشت هذه الدراسة قدرة ما يسمى بنماذج الذكاء الاصطناعي اللغوية الضخمة على إدراك المعنى، وتحديدًا ما يتعلق بمعرفة الأدوار المحورية، دون الحاجة إلى الاستعانة بخوارزميات التحليل الصرفي والنحوي وخوارزميات التحليل الدلالي ودون الاستعانة بالموارد المعجمية المطلوبة لذلك، ولضيق المساحة التي تسمح بها مثل هذه الأبحاث، اقتصرت الدراسة على اختبار قدرة نموذج الـ(ChatGPT) على تحديد المسؤول عن الحدث/ المتسبب فيه، كأحد الأدوار المحورية، بمجرد تزويده بمجموعة من الجمل وسؤاله بالطريقة التي يسأل بها الإنسان من دون الاستعانة بتلك الموارد الحاسوبية والخوارزميات المعدة لهذا العرض.

وقد كشفتْ نتائج هذه الدراسة عن ضعف أداء هذا النموذج الاصطناعي في معرفة المتسبب في الحدث بعد معرفة الكلمة الدالة على الحدث في كل جملة من الجمل المعطاة حيث تبين لنا أن متوسط أداء هذا النموذج، مقارنةً بأداء الإنسان، في مهمة تحديد الكلمة الدالة على الحدث هي (F-Socre= 0.53)، بينما حقق متوسط (F-Socre= 0.63) في معرفة المتسبب في الحدث. ومما كشفتْ عنه هذه الدراسة أن نموذج الـ(ChatGPT)، بوصفه أحد نماذج الذكاء الاصطناعي الضخمة، وإن أظهر قدرة على توليد النصوص تقارب قدرة الإنسان ما يزال غير قادر على فهم وإدراك العلاقات الدلالية داخل الجملة بصورة تقارب قدرة الإنسان على الفهم والاستنتاج؛ وبناء على ذلك، تقترح هذه الدراسة الاستفادة من هذا النموذج، نظرًا لما يوفره من جهد ووقت، في تحليل الأدوار المحورية على أن يكون ذلك بتدخل الإنسان؛ وذلك بعرض النتائج على متخصص يقوم بمراجعتها وتصحيحها، وهو ما قد يساعد في توسيم البيانات تمهيدًا لاستخدامها في مهام تدريب الآلة.

الملاحق:

الملحق1: صورة من هندسة المدخلات (Prompt Engineering)

الملحق2: صورة من بعض الجمل المدخلة

الملحق3: صورة من بعض الجمل المدخلة

الإفصاح والتصريحات:

تضارب المصالح: ليس لدى المؤلفون أي مصالح مالية أو غير مالية ذات صلة للكشف عنها. المؤلفون يعلنون عن عدم وجود أي تضارب في المصالح.

References:

(1) Tenny C, Levin B. English Verb Classes and Alternations: A Preliminary Investigation. Language (Baltim). 1995;71(1).

(2) Dowty D. Thematic Proto-Roles and Argument Selection. Language (Baltim). 1991;67(3).

(3) Palmer M, Gildea D, Xue N. Semantic role labeling. Synthesis Lectures on Human Language Technologies. 2010;3(1).

(4) Dowty D. Thematic Proto-Roles and Argument Selection. Language (Baltim). 1991;67(3).

(5) Bonial C, Hwang J, Bonn J, Conger K, Babko-Malaya O, Palmer M. English PropBank Annotation Guidelines. Technical Report, University of Colorado at Boulder. 2012;

(6) Zhao WX, Zhou K, Li J, Tang T, Wang X, Hou Y, et al. A Survey of Large Language Models. 2023;1–85.

(7) Koubaa A, Boulila W, Ghouti L, Alzahem A, Latif S. Exploring ChatGPT Capabilities and Limitations : A Critical Review of the NLP Game Changer. Preprints.org. 2023;(March):1–18.

(8) Hariri W. Unlocking the Potential of ChatGPT: A Comprehensive Exploration of its Applications, Advantages, Limitations, and Future Directions in Natural Language Processing. 2023;

(9) Haque MdA. A Brief Analysis of "ChatGPT" – A Revolutionary Tool Designed by OpenAI. EAI Endorsed Transactions on AI and Robotics. 2023;1(1).

(10) Ray PP. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Vol. 3, Internet of Things and Cyber-Physical Systems. 2023.

(11) S. OPEN DOMAIN QUESTION ANSWERING SYSTEM USING SEMANTIC ROLE LABELING. Int J Res Eng Technol. 2014;03(27).

(12) Narayanan S, Harabagiu S. Question answering based on semantic structures. In: COLING 2004 - Proceedings of the 20th International Conference on Computational Linguistics. 2004.

(13) Surdeanu M, Harabagiu S, Williams J, Aarseth P. Using predicate-argument structures for information extraction. In: Proceedings of the Annual Meeting of the Association for Computational Linguistics. 2003.

(14) Christensen J, Mausam, Soderland S, Etzioni O. An analysis of open information extraction based on semantic role labeling. In: KCAP 2011 - Proceedings of the 2011 Knowledge Capture Conference. 2011.

(15) Khan A, Salim N, Jaya Kumar Y. A framework for multi-document abstractive summarization based on semantic role labelling. Applied Soft Computing Journal. 2015;30.

(16) Zhou Q, Jiang Z, Yang F. Sentences Similarity Based on Deep Structured Semantic Model and Semantic Role Labeling. In: 2020 International Conference on Asian Language Processing, IALP 2020. 2020.

(17) Alam M, Gangemi A, Presutti V, Reforgiato Recupero D. Semantic role labeling for knowledge graph extraction from text. Progress in Artificial Intelligence. 2021;10(3):309–20.

(18) Humphreys L, Boella G, van der Torre L, Robaldo L, Di Caro L, Ghanavati S, et al. Populating legal ontologies using semantic role labeling. Artif Intell Law (Dordr). 2021;29(2).

(19) Dang HT, Palmer M. The role of semantic roles in disambiguating verb senses. In: ACL-05 - 43rd Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference. 2005.

(20) Twiefel J, Hinaut X, Wermter S. Semantic role labelling for robot instructions using echo state networks. In: ESANN 2016 - 24th European Symposium on Artificial Neural Networks. 2016.

(21) Giuglea AM, Moschitti A. Semantic role labeling via FrameNet, VerbNet and PropBank. In: COLING/ACL 2006 - 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference. 2006.

(22) Roth M, Lapata M. Context-aware Frame-Semantic Role Labeling. Trans Assoc Comput Linguist. 2015;3.

(23) Gargett A, Leung T. Building the Emirati {A}rabic {F}rame{N}et. In: Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet. 2020.

(24) Baker CF, Fillmore CJ, Lowe JB. The Berkeley FrameNet Project. In: Proceedings of the Annual Meeting of the Association for Computational Linguistics. 1998.

(25) Fillmore CJ, Johnson CR, Petruck MRL. Background to framenet. International Journal of Lexicography. 2003;16(3).

(26) Das D, Chen D, Martins AFT, Schneider N, Smith NA. Frame-Semantic Parsing. Computational Linguistics. 2014;40(1).

(27) Ellsworth M, Baker C, Petruck MRL. FrameNet and Typology. In 2021. 28. Kipper K, Korhonen A, Ryant N, Palmer M. A large-scale classification of English verbs. Lang Resour Eval. 2008;42(1).

(28) Schuler KK. VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon. Dissertation Abstracts International, B: Sciences and Engineering. 2005;66(6).

(29) Mousser J. A large coverage verb taxonomy for Abic.In: Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010. 2010.

(30) Kingsbury P, Palmer M. From TreeBank to PropBank.

(31) Palmer M, Babko-Malaya O, Bies A, Diab M,Maamouri M, Mansouri A, et al. A pilot Arabic propbank. In: Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008. 2008.

(32) Zaghouani W, Diab M, Mansouri A, Pradhan S, Palmer M. The revised Arabic PropBank. In: ACL 2010 - LAW 2010: 4th Linguistic Annotation Workshop, Proceedings. 2010.

(33) Musser J. A Large Coverage Verb Lexicon For Arabic. 2013.

(34) Gildea D, Jurafsky D. Automatic labeling of semantic roles. Computational Linguistics. 2002;28(3).

(35) Xue N, Palmer M. Calibrating features for semantic role labeling. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, EMNLP 2004 - A meeting of SIGDAT, a Special Interest Group of the ACL held in conjunction with ACL 2004. 2004.

(36) Diab M, Moschitti A, Pighin D. Semantic role labelingsystems for Arabic using Kernel Methods. ACL-08: HLT - 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference. 2008;(June):798–806.

(37) Wang S, Wang S. The Semantic Roles of Chinese Verbs of Saluting. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2023.

(38) Senator F, Boutouta H, Lakhfif A, Mediani C. Semantic Role Labeling of Arabic Emotional Text in Tweets. In: 2023 International Conference on Advances in Electronics, Control and Communication Systems, ICAECCS 2023. 2023.

(39) Mohammed MH, Al Jubouri HSA. Semantic Roles In Arabic With Reference To English. مجلة آداب الفراهيدي. 2019;

(40) Li Z, Zhao H, He S, Cai J. Syntax role for neural semantic role labeling. Computational Linguistics. 2021;47(3):529–74.

(41) Qian F, Sha L, Chang B, Liu LC, Zhang M. Syntax aware lstm model for semantic role labeling. In: EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the 2nd Workshop on Structured Prediction. 2017.

(42) Zhou J, Xu W. End-To-end learning of semantic role labeling using recurrent neural networks. In: ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference. 2015.

(43) Munir K, Zhao H, Li Z. Semi-Supervised Semantic Role Labeling with Bidirectional Language Models. ACM Transactions on Asian and Low-Resource Language Information Processing. 2023;22(6).

(44) Diab M, Fellbaum C, Alkhalifa M, Mansouri A, Elkateb S, Palmer M. Semeval 2007 task 18: Arabic semantic labeling. ACL 2007 - SemEval 2007 - Proceedings of the 4th International Workshop on Semantic Evaluations. 2007;(June):93–8.

(45) Diab M, Moschitti A, Pighin D. CUNIT: A semantic role labeling system for modern standard Arabic. ACL 2007 - SemEval 2007 - Proceedings of the 4th International Workshop on Semantic Evaluations. 2007;133–6.

(46) White J, Fu Q, Hays S, Sandborn M, Olea C, Gilbert H, et al. A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. 2023 Feb 21;