Syrian Arab Republic
Ministry of Higher Education and Scientific Research
Syrian Virtual University
Master of Applied Linguistics (MAL)

الجمهورية العربية السورية
وزارة التعليم العالي والبحث العلمي
الجامعة الافتراضية السورية
ماجستير التأهيل والتخصص في اللسانيات التطبيقية

الجامعة الافتراضية السورية
SYRIAN VIRTUAL UNIVERSITY

تأثير خوارزميات الذكاء الاصطناعي LLM المدربة في توليد المحتوى باللغة العربية

The Impact of Trained LLM AI Algorithms on Content Generation in Arabic Language

مشروع مقدم لنيل درجة الماجستير في اللسانيات التطبيقية

إشراف:

د. آلاء عيسى

إعداد الطالبة:

نتالي مؤيد البيطار

natalie_233057

(الفصل S24)

الفهارس:

فهرس المحتويات:

فهرس الجداول:

الاختصارات:

- MMLU: Massive Multitask Language Understanding

  فهم اللغة متعدد المهام على نطاق واسع

- GPQA: Graduate-Level Google-Proof Q&A Benchmark

  معيار الأسئلة والأجوبة على مستوى الدراسات العليا المعتمد على Google

الملخص:

عنوان الدراسة: تأثير خوارزميات الذكاء الاصطناعي (LLM) المدربة في توليد المحتوى باللغة العربية.

هدف الدراسة: هو التحقيق في تأثير خوارزميات الذكاء الاصطناعي المدربة في توليد المحتوى باللغة العربية، مع التركيز على تقييم أدائها وتحديد التحديات والقيود واستكشاف التطبيقات المحتملة. وعلى وجه التحديد، تهدف هذه الدراسة إلى تقييم أداء ثلاثة نماذج ذكاء اصطناعي حديثة في إنشاء محتوى عربي سليم وتوليد نصاً عربياً صحيحاً وذو صلة بالسياق. تبحث هذه الدراسة في فعالية خوارزميات الذكاء الاصطناعي في إنشاء محتوى عربي متنوع عالي الجودة.

منهج الدراسة: اعتمدت الدراسة المنهج الوصفي التحليلي الذي سمح بإجراء تقييم شامل للنماذج اللغوية (LLAMA – Gemini – GPT4) قُيم أداءها من خلال توليد مجموعة بيانات من النص العربي ومقارنتها من حيث مهام اللغة العربية (تنوع المحتوى – تحليل المشاعر – تنوع البنى النحوية التركيبية)، وحللت نتائج التجربة باستخدام الأساليب الإحصائية.

نتائج الدراسة: تظهر الدراسة مدى فعالية خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي:

1. (LLAMA 3.1–70B Instruct) تفوق في تنوع المشاعر بفضل نسب متوازنة بين المشاعر الإيجابية، السلبية، والمحايدة، (GPT–4o mini) كان الأفضل في المشاعر الإيجابية والمحايدة، لكنه افتقر للمشاعر السلبية. بينما (Gemini 1.5 F) أضعف أداءً بسبب هيمنة المحتوى الإيجابي.

2. (GPT-4o mini) تفوق بوضوح، بينما كان أداء (LLAMA 3.1-70B Instruct) و(Gemini 1.5 F) متقارباً.

3. (GPT-4o mini) تفوق في تقديم أنماط متنوعة ومتسقة، بينما افتقرت النماذج الأخرى للتنوع واعتمدت على أنماط محدودة.

4. (GPT-4o mini) النموذج الأفضل عموماً، لكنه بحاجة لتحسين في المشاعر السلبية، (LLAMA 3.1-70B Instruct) خيار جيد لتنوع المشاعر، (Gemini 1.5 F) خيار ثانوي جيد لتنوع المواضيع والبنى النحوية.

الكلمات المفتاحية: الذكاء الاصطناعي، خوارزميات الذكاء الاصطناعي، نماذج اللغة الكبيرة، توليد المحتوى، اللغة العربية.

الفصل الأول: الإطار العام للدراسة:

1.1. المقدمة:

أحدث التقدم السريع في الذكاء الاصطناعي ومعالجة اللغة الطبيعية ثورة في الطريقة التي نولد بها المحتوى ونتفاعل معه. "فقد بلغ التقدم التكنولوجي أوجه، بسبب التقدم الهائل الذي أحرزه العلم في بناء أجيال متطورة من الحاسوب. وقد كان لها التطور أن يدخل في مجالات الحياة كلها، وأن يعمل على تجديد النظر العلمي، والأساليب التي يطبقها العلماء في ميادين العلم المختلفة" (استيتية، 2008، ص527).

ظهرت نماذج اللغة الكبيرة كتكنولوجيا رئيسية في هذا المجال، مما مكن الآلات من معالجة وتوليد لغة تشبه لغة الإنسان بدقة وطلاقة غير مسبوقة. في السنوات الأخيرة، طُبقت خوارزميات الذكاء الاصطناعي المدربة على لغات مختلفة بتزايد، بما في ذلك اللغة العربية. كان تأثير هذه الخوارزميات في إنشاء المحتوى العربي كبيراً، حيث غيرت الطريقة التي يُنشئ بها المحتوى باللغة العربية ونشره واستهلاكه.

فتح دمج نماذج اللغة الكبيرة في إنشاء المحتوى العربي آفاقاً جديدة لمنشئي المحتوى، من أتمتة مهام الكتابة الروتينية إلى تعزيز عمليات الكتابة الإبداعية. جعلت نماذج اللغة الكبيرة من الممكن لمنشئي المحتوى باللغة العربية إنتاج محتوى عالي الجودة على نطاق وسرعة غير مسبوقين، بفضل القدرة على إنشاء نص دقيق وذو صلة بالسياق.

ولكن ذلك يثير العديد من المخاوف، مثل القضايا المحيطة بخصوصية البيانات، وحقوق الملكية الفكرية، والتحيزات المحتملة المتأصلة في البيانات والقرارات الخوارزمية (Geis et al., 2019; Masters, 2023). وهناك مخاوف محقة بشأن

الخسارة المحتملة للسياق الثقافي والدقة اللغوية خاصة أن اللغة العربية لغة معقدة ودقيقة ذات تراث ثقافي غني. وعلاوة على ذلك، فإن استخدامها في إنشاء المحتوى العربي يثير أيضاً أسئلة حول التأليف والملكية والمساءلة. وعلى الرغم من هذه التحديات، فإن استخدام أدوات الذكاء الاصطناعي في إنشاء المحتوى العربي لديه القدرة على تقديم الفرصة للوصول إلى محتوى باللغة العربية عالي الجودة، يمكن أن يساعد في تعزيز الحفاظ على اللغة العربية وإحيائها، وتسهيل الوصول إليها للأجيال الجديدة.

يوضح هذا البحث تأثير خوارزميات الذكاء الاصطناعي المدربة في برامج التعلم العميق في إنشاء المحتوى العربي، وفحص فوائد وتحديات وآثار هذه التكنولوجيا على اللغة العربية ومتحدثيها. سيناقش أيضاً التطبيقات المحتملة لبرامج التعلم العميق في إنشاء المحتوى العربي، وأخيراً يتناول التحديات والقيود التي يفرضها استخدام أدوات الذكاء الاصطناعي في توليد المحتوى العربي، بما في ذلك المخاوف بشأن الحساسية الثقافية والدقة اللغوية والتأليف، ومناقشة الحلول المحتملة لهذه التحديات.

1.2. مشكلة البحث:

يعتمد العصر الحالي كثيراً على تطبيقات الذكاء الاصطناعي، شمل كل المجالات والمجال اللغوي هو أحد أهمها، نتيجة لعمل الباحثة في مجال معالجة المعلومات ودراسات اللغة، لاحظت الاعتماد المتزايد على نماذج اللغة الكبيرة (LLMs) لتوليد المحتوى العربي. وفي حين بانت نماذج اللغة الكبيرة واعدة في أتمتة مهام الكتابة الروتينية وتعزيز عمليات الكتابة الإبداعية، فقد لاحظت أيضاً الافتقار إلى الحكم البشري والسياق الثقافي في المحتوى الذي يُنشئ بوساطة الذكاء الاصطناعي.

من خلال اطلاع الباحثة على العديد من الدراسات مثل (El-Shangiti et al., 2024)، (Antoun et al., 2024)، (Marie-Sainte et al., 2017) وغيرها وجدت أن نماذج اللغة الكبيرة تُستخدم لتوليد محتوى عربي بمعدل ينذر بالخطر، وغالباً دون إشراف مناسب أو مراعاة للفروق الثقافية واللغوية. وشاهدت الباحثة حالات حيث استُخدم محتوى أُنشئ بوساطة الذكاء الاصطناعي في المواد التعليمية والمقالات الإخبارية وحتى الوثائق الحكومية الرسمية، دون أي إشارة واضحة إلى ما إذا كان المحتوى قد أُنشئ بوساطة إنسان أم ذكاء اصطناعي. كان المحتوى الذي يُنشئ بوساطة هذه النماذج غير دقيق في كثير من الأحيان، لقد كان محتوى عاماً ويفتقر إلى العمق والتنوع والثراء الذي تتميز به اللغة العربية وثقافتها. وهنا يُطرح التساؤل عما إذا كنا نسيء إلى اللغة العربية باستخدام هذه التكنولوجيا، هل نضحي بالجودة من أجل السرعة والكفاءة؟

وبدأت العديد من الدول العربية في تبني هذه التكنولوجيا المتقدمة، وبدأت تُعقد العديد من المؤتمرات التي تركز على تطبيقات الذكاء الاصطناعي، أهمها مؤتمر الذكاء الاصطناعي التوليدي الذي نظمته وزارة العدل المصرية بالتعاون مع المنظمة العالمية للملكية الفكرية (WIPO) والجامعة البريطانية في مصر، الذي أوصى بالتوسع في استخدام تطبيقات الذكاء الاصطناعي في سوق العمل والمناهج الدراسية. ووضع التشريعات والقوانين لتنظيم استخدامات تطبيقات الذكاء الاصطناعي؛ تعالج مفاهيم العدالة والمساءلة والشفافية (اليوم السابع، 2024).

وأيضاً عقد المؤتمر السنوي الدولي الرابع لمركز تريندز للبحوث والاستشارات تحت عنوان «الأمن المستدام في عام (2024) وما بعده – دور الذكاء الاصطناعي» حيث دعا المشاركون في توصيات المؤتمر إلى وضع إطار عمل عالمي موحد لحوكمة الذكاء الاصطناعي، مشددين على أهمية الأخلاقيات في تطوير واستخدام هذه

التقنية، مع ضرورة ضمان استخدامها بطريقة مسؤولة وآمنة (مركز الاتحاد للأخبار، 2024).

وبعد الاطلاع على الدراسات وتوصيات المؤتمرات السابقة (وغيرها) وُجد أنه من الممكن أن يؤدي الافتقار إلى اختبار وتقييم المحتوى الذي أُنشئ بوساطة الذكاء الاصطناعي إلى عواقب قد تفقد اللغة العربية طابعها الفريد وتؤدي إلى تهالكها. من هنا جاء هذا البحث استكمالاً لهذه الدراسات واستجابة لهذه التوصيات، وإظهاراً للحاجة لاختبار وتقييم أداء أدوات الذكاء الاصطناعي في توليد المحتوى العربي وموثوقيتها وحساسيتها تجاه الثقافة العربية وتسليط الضوء على المخاطر المحتملة. فكان الهدف الأساسي من هذا البحث الإجابة على السؤال الآتي: ما تأثير خوارزميات الذكاء الاصطناعي (LLM) المدربة في توليد المحتوى باللغة العربية؟

1.3. أهمية البحث:

تأتي أهمية هذا البحث في دراسة تأثير خوارزميات الذكاء الاصطناعي لنماذج اللغة الكبيرة المدربة في توليد المحتوى العربي، واستكشاف الفوائد والتحديات والتطبيقات المحتملة لهذه التكنولوجيا. ومن خلال دراستها يمكن المشاركة في تطوير نماذج ذكاء اصطناعي أكثر تقدماً تحسن فهم اللغة العربية وتدعم ثقافتها وتعزيزها. تبرز أهمية هذه الدراسة من خلال تسليطها الضوء على النقاط الآتية:

- تحسين توليد المحتوى العربي: من خلال تحليل تأثيرات خوارزميات الذكاء الاصطناعي المدربة في توليد المحتوى العربي، يمكن للباحثين تحديد مجالات التحسين وتطوير نماذج توليد نصاً عربياً أكثر دقة وتماسكاً، والتي يمكن تطبيقها في مجالات مختلفة مثل ترجمة اللغة وتلخيص النصوص والروبوتات الدردشة.

- تحسين فهم اللغة العربية: إن دراسة تأثير خوارزميات الذكاء الاصطناعي المدربة في توليد المحتوى العربي يمكن أن تؤدي إلى فهم أعمق للغة العربية وفروقها الدقيقة وتعقيداتها، مما يؤدي في النهاية إلى تحسين أداء نماذج الذكاء الاصطناعي في معالجة وتوليد النص العربي.
- معالجة تحديات اللغة العربية: اللغة العربية لغة غنية ومعقدة من الناحية الصرفية، ولها خصائص فريدة قد تكون صعبة على نماذج الذكاء الاصطناعي التعامل معها. يمكن أن تساعد هذه الدراسة في معالجة هذه التحديات وتطوير حلول أكثر فعالية لمعالجة اللغة العربية.
- دعم الحفاظ على اللغة العربية: مع تزايد هيمنة اللغة الإنجليزية على الإنترنت، أصبحت اللغة والثقافة العربية معرضتين لخطر التهميش. ومن خلال تطوير خوارزميات الذكاء الاصطناعي الفعالة لتوليد المحتوى العربي، يمكن لهذه الدراسة أن تسهم في الحفاظ على اللغة العربية وتعزيزها.
- تطوير أبحاث معالجة اللغة الطبيعية: يمكن لهذه الدراسة أن تسهم في تطوير أبحاث معالجة اللغة الطبيعية، وخاصة في مجال اللغات ذات الموارد المنخفضة مثل العربية، والتي يمكن أن يكون لها تأثير أوسع في تطوير نماذج معالجة اللغة الطبيعية الأكثر دقة وفعالية.
- الفوائد الاقتصادية: من خلال تحسين إنشاء المحتوى العربي، يمكن أن يكون لهذه الدراسة فوائد اقتصادية، مثل تعزيز القدرة التنافسية للشركات الناطقة باللغة العربية في السوق العالمية وخلق فرص عمل جديدة في مجال معالجة اللغة الطبيعية ومعالجة اللغة العربية.
- الأهمية الثقافية: يمكن أن يكون لهذه الدراسة أهمية ثقافية من خلال الحفاظ على اللغة والثقافة العربية وتعزيزهما، وتمكين التواصل وتبادل المعلومات

بطريقة أكثر فعالية بين المجتمعات الناطقة باللغة العربية، وتسهيل إنشاء محتوى عربي يعكس ثقافة وقيم المجتمعات الناطقة باللغة العربية.

- مساعدة دراسات التحليل المقارن: من خلال دراسة تأثير خوارزميات الذكاء الاصطناعي على توليد المحتوى العربي، يمكن للباحثين إجراء تحليل مقارن مع لغات أخرى، وتحديد أوجه التشابه والاختلاف، والمساهمة في فهم أعمق لمعالجة اللغة وتوليدها.

1.4. أهداف البحث:

تهدف هذه الدراسة إلى استكشاف تأثير خوارزميات الذكاء الاصطناعي المدربة في إنتاج المحتوى باللغة العربية، مع التركيز على تقييم أدائها وتحديد التحديات والقيود واستكشاف التطبيقات المحتملة. وتوضح الأهداف الآتية نطاق واتجاه هذا البحث، الذي يسعى إلى المساهمة في تطوير نماذج الذكاء الاصطناعي أكثر تقدماً وفهماً للغة العربية:

- التحقق من فعالية خوارزميات الذكاء الاصطناعي المدربة في توليد محتوى عربي متنوع: نهدف إلى تقييم أداء خوارزميات الذكاء الاصطناعي في توليد نصاً عربياً صحيحاً.
- تحديد التحديات والقيود التي تواجه خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي: نهدف إلى التحقيق في الصعوبات والعقبات التي تواجهها خوارزميات الذكاء الاصطناعي في توليد محتوى عربي عالي الجودة، بما في ذلك القضايا المتعلقة بتعقيد اللغة وجودة البيانات والفروق الثقافية.
- توضيح كيفية عمل أداء خوارزميات الذكاء الاصطناعي المختلفة في توليد المحتوى العربي: يسعى هذا الهدف إلى تقييم ومقارنة أداء خوارزميات الذكاء الاصطناعي المختلفة في توليد المحتوى العربي.

- دراسة تأثير حجم بيانات التدريب وجودتها في أداء خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي.

1.5. أسئلة البحث:

- ما مدى فعالية خوارزميات الذكاء الاصطناعي المدربة في توليد محتوى عربي متنوع؟
- ما التحديات والقيود التي تواجه خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي؟
- كيف تعمل خوارزميات الذكاء الاصطناعي المختلفة في توليد المحتوى العربي، وأي خوارزمية هي الأكثر فعالية؟
- ما تأثير حجم بيانات التدريب وجودتها في أداء خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي؟

1.6. منهج البحث:

سيكون نهج هذا البحث المنهج الوصفي التحليلي، لمقارنة أداء خوارزميات الذكاء الاصطناعي في نماذج (LLM) في مهام إنشاء المحتوى العربي. وتقييم أداء كل خوارزمية ذكاء اصطناعي في نماذج (LLM) على مجموعة من المهام والمقاييس المحددة مسبقاً.

ويعرف الباحثون المنهج الوصفي التحليلي اصطلاحاً بأنه "أسلوب من أساليب التحليل المركز على معلومات كافية ودقيقة عن ظاهرة أوموضوع محدد، أو فترة أو فترات زمنية معلومة، وذلك من أجل الحصول على نتائج علمية، ثم تفسيرها بطريقة موضوعية، بما ينسجم مع المعطيات الفعلية للظاهرة" (دويدري، 2000، ص183).

1.6.1. إجراءات الدراسة:

1) اختيار نماذج (LLM): تُختار مجموعة من نماذج (LLM) المدربة مسبقاً لتقييمها، مع مراعاة عوامل مثل البنية وبيانات التدريب والأداء على لغات أخرى.
2) جمع البيانات: إنشاء مجموعة بيانات واسعة النطاق من الجمل في اللغة العربية باستخدام النماذج المدربة (LLM) المسبق اختيارها، تضم أنواعاً وأنماطاً وموضوعات مختلفة. بحيث تُحدد عينة تمثيلية، مع ضمان التنوع والتوازن من حيث طول الجملة وتعقيدها وأسلوبها.
3) إجراء التحليل والنتائج: يُقيم النص الناتج، باستخدام إطار تقييم موحد وتحليل النتائج وتحديد الأنماط.

1.7. حدود البحث:

فيما يأتي بعض القيود البحثية المعتمدة في هذه الدراسة، من خلال الاعتراف بهذه القيود يمكن توفير فهم أكثر دقة لنتائج الدراسة وتسليط الضوء على مجالات البحث المستقبلية:

1.7.1. الحدود الزمانية:

أجريت الدراسة خلال الفصل الدراسي S24 من شهر أيلول إلى شهر كانون الأول من العام (2024).

1.7.2. الحدود الموضوعية:

ستفحص الدراسة فقط تأثير خوارزميات الذكاء الاصطناعي في بعض النماذج المحددة المدربة والتي لها حجماً محدداً من بيانات التدريب. وتركز الدراسة على نماذج (LLM) المدربة التي أُصدرت في آخر عامين من تاريخ إجراء هذه الدراسة (2023–

2024) ويجب الأخذ بعين الاعتبار أنه لا يمكن إسقاط التطورات في خوارزميات الذكاء الاصطناعي والتي ربما تُصدر أو تُحدث بعد فترة جمع بيانات الدراسة حيث يعد هذا المجال سريع التطور بدرجة أكبر.

يركز هذا البحث على توليد المحتوى المكتوب ولن يدرس أشكالاً أخرى من توليد المحتوى مثل توليد الصور أو الفيديو. سيكون هذا المحتوى باللغة العربية الفصحى الحديثة ولن يدرس اللهجات العربية الأخرى، مثل العربية المصرية أو العربية الشامية.

1.8. مصطلحات البحث:

- **خوارزميات الذكاء الاصطناعي:**

تسبق الخوارزميات الذكاء الاصطناعي وقد استُخدمت على نطاق واسع خارج هذا المجال، لكن في سياق الذكاء الاصطناعي، الخوارزمية هي إجراء محدد جيداً يأخذ بعض المدخلات ويعالجها وينتج مخرجات مقابلة.

التعريف الاصطلاحي: مصطلح "الخوارزمية" مشتق من اسم عالم الرياضيات الفارسي في القرن التاسع محمد بن موسى الخوارزمي ويشير إلى تعليمات محددة لحل مشكلة أو إجراء حساب (Sheikh et al., 2023). وعُرفت خوارزمية الذكاء الاصطناعي بأنها "مجموعة من التعليمات التي تمكن الكمبيوتر من أداء مهمة محددة تتطلب عادةً ذكاءً بشرياً، مثل التعلم وحل المشكلات واتخاذ القرار والإدراك أو فهم اللغة" (Russell, Norvig, 1995).

وتعرف الباحثة خوارزميات الذكاء الاصطناعي بأنها خوارزميات صممت لمعالجة وتحليل كميات هائلة من البيانات، وتحديد الأنماط، واتخاذ القرارات أو التنبؤات بناءً على تلك البيانات، لتمكين الآلات من أداء المهام التي تتطلب عادةً الذكاء البشري.

وتشكل خوارزميات الذكاء الاصطناعي العمود الفقري لأنظمة الذكاء الاصطناعي، مما يسمح لها بالتعلم من الخبرة، والتكيف مع المواقف الجديدة، وتحسين أدائها بمرور الوقت.

- **نماذج اللغة الكبيرة المدربة Pre-Trained Large Language Models:**

التعريف الاصطلاحي: تشير نماذج اللغة الكبيرة (LLMs) إلى نماذج اللغة المحولة التي تحتوي على مئات المليارات (أو أكثر) من المعلمات، والتي تُدرب على بيانات نصية ضخمة. تُظهر نماذج اللغة الكبيرة قدرات قوية لفهم اللغة الطبيعية وحل المهام المعقدة (عبر إنشاء النص) (Zhao et al., 2023) وتُدرب هذه النماذج مسبقاً في مجموعة كبيرة من النصوص ثم ضبط هذه النماذج في مهام لاحقة مختلفة لتحقيق نتائج متطورة (Li et al., 2021).

وتعرف الباحثة نماذج اللغة الكبيرة المدربة إجرائياً بأنها المحاكاة الأكثر إقناعاً للغة البشرية والتي أُنشئت بوساطة الحاسوب حتى الآن، والتي تُدرب على مجموعة كبيرة من بيانات النص لتعلم أنماط وهياكل اللغة. صُممت نماذج اللغة المدربة مسبقاً لالتقاط الخصائص الإحصائية للغة، ويمكن ضبطها بدقة لمهام معالجة اللغة الطبيعية المحددة. فهي يمكنها التعرف إلى اللغات البشرية وتلخيصها وترجمتها والتنبؤ بها وتوليدها على أساس مجموعات بيانات نصية كبيرة جداً، نظراً لأن اللغة التي تولدها نماذج اللغة الكبيرة ستكون أكثر تعقيداً وشبهاً بالتي يولدها الإنسان من أسلافها.

- **توليد المحتوى:**

أدى النمو في تطبيق المحتوى المولد بوساطة الذكاء الاصطناعي إلى تغيير الطريقة التي تولد بها الشركات كميات هائلة من النصوص والصور ومقاطع الفيديو بسرعة ودقة (Davenport et al., 2020).

لقد شهد المحتوى المولد بوساطة الذكاء الاصطناعي استخداماً واسع النطاق بسبب فعاليته في مهام معينة، ومع ذلك فإن هذا الاستخدام ليس خالياً من العيوب. إن إدراك كيفية تأثير المواد التي تُنشئ بوساطة الذكاء الاصطناعي في التفاعل يمكن أن يساعد في إنشاء محتوى أفضل جودة (Bilgihan et al., 2016).

التعريف الاصطلاحي: يشير مصطلح إنشاء المحتوى في الذكاء الاصطناعي إلى عملية إنشاء النصوص أو الصور أو مقاطع الفيديو أو أشكال أخرى من الوسائط تلقائياً باستخدام خوارزميات الذكاء الاصطناعي، والتي تتضمن تدريب نماذج التعلم الآلي على مجموعات بيانات كبيرة من المحتوى الموجود لتعلم الأنماط والهياكل، ثم استخدام هذه النماذج لإنشاء محتوى جديد مشابه في الأسلوب والجودة لبيانات التدريب (Goldberg, 2017).

وتعرف الباحثة إنشاء المحتوى إجرائياً بأنه عملية إنشاء محتوى جديد مثل النص، باستخدام الخوارزميات الحاسوبية. في سياق (LLM) يشير إنشاء المحتوى إلى قدرة النموذج على إنشاء نص متماسك ومحدد للسياق. يمكن أن يشمل هذا إنشاء نص لتطبيقات مختلفة، مثل برامج الدردشة الآلية، وترجمة اللغة، وتلخيص النص.

- **اللغة العربية:**

انتشرت اللغة العربية على نطاق واسع بعد ظهور الإسلام، على الرغم من أنها كانت موجودة قبل الدين بقرون. تاريخياً تتجذر اللغة العربية في اللغة العربية

الفصحى، والتي استخدمت كلغة أصلية للشعوب العربية منذ عام (600) بعد الميلاد. ومع ذلك، على مر القرون، تطورت اللغة وبسطت لإنشاء ما يُعرف بالعربية الفصحى الحديثة وأصبح لكل منطقة لهجة من اللغة العربية يتحدث بها في المجتمع (Farghaly & Shaalan, 2009).

التعريف الاصطلاحي: اللغة العربية هي لغة أفروآسيوية تطورت في الشرق الأوسط. ويتحدث بها أكثر من (250) مليون شخص في جميع أنحاء العالم (Al-Ajlan et al., 2008). وعرفت أيضاً بأنها لغة إنسانية حية، لها نظامها الصوتي والصرفي والنحوي والتركيبي، كما لألفاظها دلالاتها الخاصة بها. وقد رأى العلماء أن كل خروج عن هذا النظام اللغوي المتكامل يعد لحناً، سواء أكان هذا الخروج بخلط الكلام بلغة أخرى، أم باستعمال اللفظة في غير موضعها، أم في مخالفة أي عنصر أساسي من عناصر كيانها اللغوي التي يميزها عن غيرها من اللغات الإنسانية (المقدسي، 1959).

وتعرف الباحثة اللغة العربية إجرائياً بأنها لغة فريدة ومعقدة لها قواعدها النحوية وتراكيبها اللغوية وأبجديتها الخاصة. تتمتع بموسوعة مفردات غنية، يتم دمجها بطرق مختلفة لتكوين كلمات جديدة، حيث يسمح نظامها بالتعبير عن الأفكار والمفاهيم المعقدة بدقة ووضوح. واللغة العربية لغة موحدة إلى حد كبير، ولها تراث أدبي وثقافي مشترك يتجاوز الحدود الإقليمية والوطنية.

- **توليد المحتوى باللغة العربية:**

يجب أن تتعامل تطبيقات معالجة اللغة الطبيعية العربية مع العديد من المشكلات المعقدة ذات الصلة بطبيعة وبنية اللغة العربية. وباعتبارها لغة طبيعية، فإن اللغة العربية أيضاً فريدة من نوعها من حيث تاريخها وطبيعتها الثنائية اللسانية وبنيتها

الداخلية وارتباطها الذي لا ينفصم بالإسلام والثقافة والهوية العربية. ومن المؤكد أن أي نظام معالجة طبيعية للغة العربية لا يأخذ في الاعتبار السمات المحددة للغة العربية غير كافٍ (Shaalan, 2005).

التعريف الاصطلاحي: إن توليد المحتوى باللغة العربية مجالاً بحثياً مثيراً للاهتمام. فهو يتضمن تطوير تقنيات وأدوات باستخدام اللغة العربية. وقد أُنشئ العديد من الأنظمة الحالية لتطبيقات مختلفة مثل الترجمة الآلية، واسترجاع المعلومات واستخراجها، والتوطين، وأنظمة استرجاع المعلومات متعددة اللغات (Larkey et al., 2002) (Habash, 2007).

تعرف الباحثة توليد المحتوى باللغة العربية إجرائياً بأنه عملية إنشاء محتوى جديد باللغة العربية باستخدام خوارزمية حاسوبية. ويمكن أن يشمل ذلك إنشاء نص لتطبيقات مختلفة، مثل برامج الدردشة الآلية، وترجمة اللغة، وتلخيص النص. يتضمن إنشاء المحتوى العربي باستخدام (LLM) تدريب النموذج على مجموعة كبيرة من النصوص العربية، ثم استخدام النموذج لإنشاء نص بناءً على موجه أو موضوع.

الفصل الثاني: بعض الدراسات السابقة:

1. عنوان الدراسة:

A Review of Artificial Intelligence Algorithms in Document Classification (Bilski, 2011)

يقدم الباحث نظرة عامة على الحالة الحالية لخوارزميات الذكاء الاصطناعي في تصنيف المستندات. تبدأ الدراسة بتقديم مفهوم تصنيف المستندات وأهميته في مجالات مختلفة، مثل تصنيف النصوص وتحليل المشاعر واسترجاع المعلومات. يسلط الباحث الضوء على التحديات المرتبطة بتصنيف المستندات، بما في ذلك الأبعاد العالية لبيانات النص والضوضاء والغموض والحاجة إلى خوارزميات فعالة ودقيقة. تهدف الدراسة إلى مراجعة خوارزميات الذكاء الاصطناعي الحالية المستخدمة في تصنيف المستندات، ومناقشة نقاط قوتها وضعفها وتطبيقاتها.

تستعرض الدراسة خوارزميات التعلم الخاضع للإشراف وأنواعها، ويناقش الباحث مزايا وعيوب كل خوارزمية، بما في ذلك أدائها على مجموعات البيانات والمهام المختلفة. وتغطي الدراسة أيضاً استخدام أساليب المجموعة، مثل التعبئة والتعزيز، وتطبيقاتها في تصنيف المستندات. تتناول الدراسة بعد ذلك خوارزميات التعلم العميق، بما في ذلك الشبكات العصبية التلافيفية، والشبكات العصبية المتكررة، وشبكات الذاكرة الطويلة والقصيرة المدى. ويناقش الباحث تطبيقات هذه الخوارزميات في تصنيف المستندات، بما في ذلك تصنيف النصوص، وتحليل المشاعر، ونمذجة الموضوعات. ثم استعرضت الدراسة خوارزميات التعلم غير الخاضع للإشراف، بما في ذلك تقنيات التجميع وتقليل الأبعاد. يناقش الباحث استخدام هذه الخوارزميات في تصنيف الوثائق، بما في ذلك تحديد المواضيع والموضوعات.

تختتم الدراسة بتسليط الضوء على أهمية خوارزميات الذكاء الاصطناعي في تصنيف الوثائق وتطبيقاتها في مجالات مختلفة. ويؤكد الباحث على الحاجة إلى مزيد من البحث في هذا المجال، بما في ذلك تطوير خوارزميات أكثر كفاءة ودقة. تقدم الدراسة نظرة شاملة على الحالة الحالية لخوارزميات الذكاء الاصطناعي في تصنيف الوثائق وتطبيقاتها.

2. **عنوان الدراسة:**

Arabic Natural Language Processing and Machine Learning-based Systems (Marie-Sainte et al., 2017)

تقدم الدراسة نظرة عامة على معالجة اللغة الطبيعية العربية وتطبيقاتها في أنظمة التعلم الآلي. يبدأ المؤلفون بخلفية موجزة عن اللغة العربية وتاريخها وخصائصها الفريدة. تناقش الدراسة اللهجات المختلفة للغة العربية، ونظام الكتابة، والميزات اللغوية التي تجعل اللغة العربية لغة صعبة المعالجة. كما يستعرض المؤلفون الحالة الحالية لمعالجة اللغة الطبيعية العربية، بما في ذلك الموارد والأدوات والتقنيات المتاحة.

تقدم الدراسة مراجعة شاملة للطرائق القائمة على التعلم الآلي لمهام معالجة اللغة الطبيعية باللغة العربية، بما في ذلك تصنيف النصوص وتحليل المشاعر وترجمة اللغة. كما تستعرض الدراسة تقنيات استخراج الميزات المختلفة المستخدمة للنص العربي، وقدمت نتائج العديد من التجارب التي أجريت على مهام معالجة اللغة الطبيعية باللغة العربية باستخدام الطرائق القائمة على التعلم الذكي.

وخلصت الدراسة إلى أن معالجة اللغة الطبيعية باللغة العربية مهمة صعبة بسبب تعقيد اللغة العربية والتوافر المحدود للموارد. ومع ذلك، تظهر الدراسة أن الأساليب

القائمة على التعلم الآلي يمكن أن تحقق أداءً متطوراً في العديد من مهام معالجة اللغة. وتقدم الدراسة توصيات عدة للبحوث المستقبلية في هذا المجال، بما في ذلك تطوير أنظمة معالجة أكثر دقة وكفاءة باستخدام نماذج التعلم العميق وإنشاء مجموعات بيانات أكبر وأكثر تنوعاً وتطوير مقاييس تقييم أكثر قوة ودقة.

3. **عنوان الدراسة**

Evaluating ChatGPT and Bard AI on Arabic Sentiment Analysis (Al-Thubaity et al., 2023)

تقدم الدراسة تقييماً لأداء نموذجين شائعين للذكاء الاصطناعي (ChatGPT) و(Bard AI)، في مهام تحليل المشاعر. يهدف الباحثون إلى تقييم فعالية هذه النماذج في اكتشاف قطبية المشاعر (إيجابية أو سلبية أو محايدة) في النص العربي. اختار الباحثون هذه النماذج بسبب شعبيتها وتنوعها في مهام معالجة اللغة الطبيعية.

لتقييم أداء (ChatGPT) و(Bard AI)، استخدم الباحثون مجموعة بيانات مكونة من 10000 عينة نصية عربية، مع شرح علامات المشاعر (إيجابية أو سلبية أو محايدة). قاموا بضبط كلا النموذجين على مجموعة البيانات وتقييم أدائهما باستخدام مقاييس مثل الدقة والقياسية والتذكر ودرجة تظهر النتائج أن كلاً من (ChatGPT) و(Bard AI) يحققان دقة عالية في تحليل المشاعر، ولكن مع بعض الاختلافات. كما تقارن الدراسة أداء (ChatGPT) و(Bard AI) مع نماذج أخرى متطورة في تحليل المشاعر. وتُظهر النتائج أن كلاً من (ChatGPT) و(Bard AI) يتفوقان على النماذج الأخرى من حيث الدقة. ومع ذلك، تشير الدراسة أيضاً إلى أن مجموعة البيانات المستخدمة محدودة، وهناك حاجة إلى مزيد من التقييم على مجموعات بيانات أكبر وأكثر تنوعاً لتأكيد النتائج. وتخلص الدراسة إلى أن (ChatGPT) و(Bard AI)

فعالان في تحليل المشاعر العربية، لكن أداء (ChatGPT) أفضل قليلاً. ويسلط الباحثون الضوء على أهمية ضبط هذه النماذج على مجموعات البيانات العربية لتحقيق الأداء الأمثل.

4. **عنوان الدراسة:**

Arabic Automatic Story Generation with Large Language Models (El-Shangiti et al., 2024)

تبحث الدراسة في تطبيق نماذج لغوية كبيرة في إنشاء القصص العربية تلقائياً. ويهدف الباحثون إلى استكشاف إمكانات نماذج اللغة الكبيرة في إنشاء قصص عربية عالية الجودة. تفرض اللغة العربية تحديات فريدة لمهام معالجة اللغة الطبيعية (NLP) بسبب امتلاكها تركيباً نحوياً معقداً، ومورفولوجيا غنية، وتوافر محدود لمجموعات البيانات المصنفة.

استخدم الباحثون نموذجاً لغوياً قائماً على محول مصمماً خصيصاً للغة العربية لإنشاء القصص. ضُبط النموذج على مجموعة بيانات كبيرة من النصوص العربية، بما في ذلك القصص والمقالات والكتب. يستخدم المؤلفون مزيجاً من نمذجة اللغة ومهام التنبؤ بالجملة التالية لالتقاط السياق المحلي والعالمي. قُيم أداء النموذج باستخدام مقاييس مختلفة، كما جُربت أيضاً معايير فرعية مختلفة لتحسين أداء النموذج.

تقدم الدراسة نتائج واعدة، حيث يولد النموذج قصصاً جيدة وقابلة للمقارنة بالقصص التي كتبها البشر. يبلغ الباحثون عن تحسينات كبيرة في أداء النموذج عند استخدام مجموعة بيانات كبيرة وضبط النموذج على نوع معين من النص. تظهر

النتائج أيضاً أن النموذج قادر على التقاط الفروق الدقيقة للغة العربية، بما في ذلك تعقيداتها الصرفية والنحوية.

توضح الدراسة إمكانات نماذج اللغة الكبيرة في توليد القصص العربية تلقائياً. يستنتج الباحثون أنه يمكن استخدام نهجهم لتوليد قصص عالية الجودة باللغة العربية، والتي يمكن أن تكون مفيدة لتطبيقات مختلفة مثل إنشاء المحتوى وتعلم اللغة والترفيه. ويسلط الباحثون الضوء أيضاً على القيود التي تواجه دراستهم ويقترحون اتجاهات بحثية مستقبلية، بما في ذلك استكشاف أنواع أخرى من النصوص، واستخدام المدخلات متعددة الوسائط، ودمج ردود الفعل البشرية لتحسين أداء النموذج.

5. **عنوان الدراسة:**

Evaluating the Effectiveness of Pre-trained Language Models in Predicting the Helpfulness of Online Product Reviews (Boluki et al., 2024)

توضح الدراسة استخدام نماذج اللغة المدربة مسبقاً في التنبؤ بمدى فائدة مراجعات المنتجات عبر الإنترنت. جمع المؤلفون مجموعة بيانات من مراجعات المنتجات من منصة تجارة إلكترونية شهيرة وقاموا بمعالجة البيانات مسبقاً عن طريق إزالة الأحرف الخاصة، وتحويل النص إلى أحرف صغيرة، ورمزية النص. اختار المؤلفون العديد من نماذج اللغة المدربة مسبقاً، بما في ذلك (BERT) و(RoBERTa) و(XLNet)، وقارنوا أداءها بنماذج التعلم الآلي التقليدية. أظهرت النتائج أن النماذج المدربة تفوقت على النماذج التقليدية في التنبؤ بمدى فائدة مراجعات المنتجات عبر الإنترنت. على وجه التحديد، حقق (BERT) و(RoBERTa) أعلى

درجات الدقة. ووجدت الدراسة أيضاً أن النماذج المدربة فعالة في التقاط مشاعر المراجعة والجودة، وهو جانب حاسم للتنبؤ بمدى فائدتها.

إن نتائج الدراسة لها آثار كبيرة على تطوير منصات التجارة الإلكترونية. من خلال استخدام نماذج دورة حياة المنتج للتنبؤ بمدى فائدة المراجعات، يمكن لمنصات التجارة الإلكترونية تحسين جودة مراجعات المنتجات وتعزيز تجربة العملاء عموماً. يمكن أن يؤدي هذا إلى زيادة رضا العملاء وولائهم، فضلاً عن تحسين المبيعات والإيرادات.

6. عنوان الدراسة:

From Text to Source: Results in Detecting Large Language Model-Generated Content (Antoun et al., 2024)

يقدم البحث دراسة المحتوى الناتج عن نماذج اللغة، والذي أصبح قضية ملحة في مجال معالجة اللغة الطبيعية. ومع تطور نماذج اللغة الكبيرة، أصبح من الصعب التمييز بين النص المكتوب بوساطة الإنسان والنص الناتج عن الذكاء الاصطناعي. يؤكد المؤلفون أن دراسة المحتوى الناتج عن نموذج اللغة الكبيرة أمر بالغ الأهمية للحفاظ على سلامة المعلومات عبر الإنترنت ومنع انتشار المعلومات المضللة. كما يقترحون نهجاً جديداً لدراسة المحتوى الناتج عن نموذج اللغة الكبيرة من خلال تحليل السمات اللغوية للنص. ويقوم المؤلفون بتقييم طريقتهم على مجموعة كبيرة من النصوص المكتوبة بوساطة الإنسان والنصوص التي أُنشئت بوساطة الذكاء الاصطناعي.

يستخدمون مجموعة من خوارزميات التعلم الذكي لتدريب "المصنف" للكشف عن المحتوى الذي أُنشئ بوساطة (LLM)، ويحققون نتائج واعدة. يتمكن "المصنف" من التمييز بدقة بين النص المكتوب بوساطة الإنسان والنص الذي أُنشئ بوساطة

(LLM) بدقة عالية، حتى عندما يكون النص الذي أُنشئ بوساطة (LLM) عالي الجودة.

أحد النتائج الرئيسية للدراسة هو أن النص الذي أُنشئ بوساطة (LLM) يميل إلى أن يكون له "أسلوب" مميز يختلف عن النص المكتوب بوساطة الإنسان. على سبيل المثال، يميل النص الذي أُنشئ بوساطة (LLM) إلى أن يكون أكثر رسمية، مع أخطاء نحوية أقل ونبرة أكثر اتساقاً. ومع ذلك، وجد المؤلفون أيضاً أن النصوص التي أُنشئت بوساطة (LLM) غالباً ما تفتقر إلى الفروق الدقيقة والإبداع الموجود في النصوص المكتوبة بوساطة الإنسان، ويمكن التعرف عليها بسهولة من خلال تحليل سماتها اللغوية.

يناقش المؤلفون الآثار المترتبة على نتائجهم على اكتشاف المحتوى الذي أُنشئ بوساطة (LLM)، ويسلطون الضوء على أهمية تطوير أساليب أكثر تطوراً للكشف عن النصوص التي أُنشئت بوساطة الذكاء الاصطناعي. كما يلاحظون أن دراستهم لها آثار على تطوير أنظمة ذكاء اصطناعي أكثر شفافية وقابلية للتفسير، يمكنها توليد نصوص أكثر توافقاً مع التفضيلات البشرية وإنشاء تدابير مضادة أكثر فعالية ضد انتشار المعلومات المضللة (LLM)، وزيادة ثقة المستخدم في المحتوى الذي يُنشئ بوساطة الذكاء الاصطناعي.

تعقيب على الدراسات السابقة:

ألهمت هذه الدراسات الباحثة ودفعتها لإجراء بحث علمي حول تأثير خوارزميات الذكاء الاصطناعي (LLM) المدربة في توليد المحتوى باللغة العربية. وضحت بعض هذه الدراسات السابقة تأثير خوارزميات نماذج اللغة في إنشاء المحتوى بلغات مختلفة ومنها العربية، مما أرسى الأساس لمزيد من البحث في هذا المجال. لكن توجه معظمهم لتعديل هذه النماذج بأساليب عدة في هدف أداء مهام محددة، لكنهم واجهوا تحديات بسبب تعقيد اللغة العربية. أظهرت هذه الدراسات نتائج واعدة في استخدام مناهج التعلم العميق لنمذجة اللغة العربية، لكن هذه الجهود كانت محدودة إلى حد كبير بمجموعات البيانات الصغيرة والمهام المحددة.

ومع ذلك يختلف البحث الحالي عن الدراسات السابقة في طرائق رئيسة عدة. أولاً، بينما ركزت الدراسات السابقة على مهام محددة، يهدف البحث إلى التحقيق في تأثير (LLM) في إنشاء المحتوى العربي باستخدام إطار تقييم أكثر شمولاً. بالإضافة إلى ذلك، بينما ركزت الدراسات السابقة في المقام الأول على مهام الترجمة الآلية أو تصنيف النصوص، يهدف هذا البحث إلى تحليل أثر استخدام نماذج اللغة في إنشاء نصاً عربياً متنوعاً وعالي الجودة لمهام إنشاء المحتوى المختلفة. علاوة على ذلك، ستبحث الدراسة أيضاً في تأثير هياكل اللغة واستراتيجيات التدريب المختلفة على إنشاء المحتوى العربي، والذي لم يتوضح بدقة في الدراسات السابقة.

الفصل الثالث: الإطار النظري للدراسة:

3.1. اللسانيات الحاسوبية:

3.1.1. المفهوم والتعريف:

تعد اللسانيات الحاسوبية من أهم فروع اللسانيات التطبيقية التي انصهرت فيها اللسانيات وعلم الحاسوب. هي مجال متعدد التخصصات يسعى إلى فهم العلاقة المعقدة بين اللغة البشرية والحوسبة. يستعين هذا المجال بالرؤى المستمدة من اللغويات وعلوم الحاسوب والذكاء الاصطناعي والعلوم المعرفية لتطوير فهم أعمق لبنية اللغة البشرية.

تعرف اللسانيات الحاسوبية بأنها "دراسة علمية للغة الطبيعية من منظور حاسوبي، وهذه الدراسة لا يمكن أن تتم إلا ببناء برامج حاسوبية لأنظمة اللغات البشرية من خلال تقييس ومحاكاة نظام عمل الدماغ البشري لنظم عمل الحاسب الآلي" (مهديوي، 2008).

تعددت التسميات التي تتناول هذا العلم فأطلق عليه اللسانيات الحاسوبية، اللغويات الحاسوبية، علم اللغة الحاسبي، لكن ما يزال تعريفه غير مستقر نتيجة اختلاف المنطلقات والمرجعيات والترجمات. وعرفها أحد الباحثين أيضاً بأنها "علم يعنى باستخدام الحاسوب وتطبيق مناهج العلوم المعتمدة عليه في دراسة اللغة، ولاسيما في الترجمة الآلية، وتمييز الكلام والذكاء الاصطناعي، أي العمليات التي تقوم بها الآلة بعد تلقينها المعلومات في حقل معين" (بعلبكي، 1990).

من هنا تستنتج الباحثة بأن اللسانيات الحاسوبية هي علم يهتم بدراسة اللغة باستخدام الحاسوب وتطبيق المناهج التي تعتمد عليه للكشف عن الأنماط والهياكل الخفية التي تكمن وراء اللغة البشرية. يهتم علم اللغة الحاسوبية بتطوير النماذج

الحاسوبية التي يمكنها التقاط تعقيدات اللغة البشرية، وفهم كيفية استخدام اللغة في العالم الحقيقي، وكيف تشكلها السياقات الاجتماعية والثقافية والتاريخية التي تُستخدم فيها، من أجل أهداف محاكاتها واستخدامها لأغراض التحليل والترجمة وتوليد المحتوى.

3.1.2. لمحة تاريخية عن تطور اللسانيات الحاسوبية:

إن علم اللغويات الحاسوبية ليس علماً جديداً. فمنذ خمسينيات القرن العشرين جرت محاولات لاستخدام أجهزة الحاسوب لمعالجة وتحليل اللغة البشرية وقد ركزت هذه المحاولات أساساً على الترجمة الآلية، وبسبب الوضع السياسي في ذلك الوقت، ركزت شبه حصرياً على الترجمة من اللغة الروسية إلى اللغة الإنجليزية. وقد تأثر هذا العمل المبكر بتطوير أول أجهزة الحاسوب وظهور الذكاء الاصطناعي كمجال بحثي. أدى تقاطع اللغويات وعلوم الحاسوب والذكاء الاصطناعي إلى إنشاء مجال جديد يركز على تطوير الخوارزميات والنماذج الإحصائية لمعالجة وفهم اللغة البشرية.

وفي الستينيات من القرن العشرين، بدأ الباحثون في استكشاف إمكانيات اللغويات الحاسوبية، وتطوير خوارزميات مبكرة ونماذج إحصائية لمعالجة اللغة. ركز هذا العمل المبكر على الترجمة الآلية وتحليل النصوص وتحليل اللغة. طُورت برامج معالجة اللغة الطبيعية الأولى خلال هذه الفترة، بما في ذلك تجارب أثبتت جدوى الترجمة الآلية.

شهدت فترة السبعينيات والثمانينيات من القرن العشرين تأسيس اللغويات الحاسوبية كمجال متميز. وقد قدم باحثون مثل نعوم تشومسكي وجون مكارثي ومارفين مينسكي مساهمات كبيرة في تطوير معالجة اللغة الطبيعية. وقد أرسى نشر كتاب

تشومسكي "الهياكل النحوية" الأساس لتطوير نظرية اللغة الرسمية، والتي لا تزال حجر الزاوية في اللغويات الحاسوبية اليوم.

في تسعينيات القرن العشرين والعقد الأول من القرن الحادي والعشرين، خضعت اللغويات الحاسوبية لتحول كبير مع ظهور تقنيات التعلم الآلي والتعلم العميق. وقد مكنت هذه التطورات من تطوير أنظمة معالجة اللغة الطبيعية أكثر دقة وقوة، ومهدت الطريق لاعتماد معالجة اللغة الطبيعية على نطاق واسع في الصناعة والأوساط الأكاديمية. واستمر المجال في التطور، مع تطوير خوارزميات جديدة ونماذج إحصائية وتقنيات التعلم الآلي التي من شأنها أن تشكل مسار اللغويات الحاسوبية في العقود القادمة (Johri et al., 2020, p2–4).

3.2. أهمية اللسانيات الحاسوبية:

تلعب اللغويات الحاسوبية دوراً حيوياً في عصرنا الرقمي الحالي، حيث تُعَد اللغة الوسيلة الأساسية للتفاعل بين الإنسان والحاسوب. تكمن أهمية اللغويات الحاسوبية في قدرتها على تمكين أجهزة الحاسوب من فهم اللغة البشرية وتفسيرها وتوليدها، وبالتالي تسهيل التواصل الفعال بين البشر والآلات. ويُذكر من الفوائد التي تقدمها الآتي:

- تخزين أكبر كم من المواد اللغوية، وما يتعلق بها من شروح في أقراص بسيطة، صغيرة الحجم وتنظيمها، لتسهيل عملية الاستفادة منها (داود، 2001).
- تمكين تعلم اللغة: تسهيل تطوير أدوات وموارد تعلم اللغة، مما يجعل من السهل على الأشخاص تعلم اللغة الأم ولغات جديدة أخرى.

- الوصول الفعال للمعلومات من خلال السرعة في التنقل بينها وفلترتها ومعالجتها; إذ إن تكنولوجيا اللغات الإنسانية لإدارة المحتوى شرط ضروري لتحويل ثروة المعلومات الرقمية إلى معرفة جماعية (بشار، 2020).
- تسهيل تحليل المشاعر: يتيح تحليل بيانات النصوص لفهم الرأي العام والمشاعر والعواطف.
- تحسين إنشاء المحتوى: يتيح تطوير الأدوات التي تساعد في إنشاء المحتوى، مثل مولدات اللغة وملخصات النصوص.
- تسهم اللسانيات الحاسوبية في تسهيل عمل المعجمي من خلال الإحصاء والتصنيف والترتيب والمقارنة والتلخيص الآلي، وتسهيل الترجمة بعمليتين فهم النص الأصلي، والتعبير عن المحتوى والأسلوب بلغة أخرى (بشار، 2020، ص8).

3.3. مجالات اللسانيات الحاسوبية:

سعى علماء اللغة حديثاً لمعالجة الظواهر اللغوية حاسوبياً، ومن خلال ذلك انتشرت تطبيقات اللغويات الحاسوبية في العديد من المجالات، وتغيرت الطريقة التي نتفاعل بها مع اللغة والتكنولوجيا، ويُذكر منها:

- التعلم اللغوي بمساعدة الحاسوب
- الإحصاء اللغوي (السباعي، 2022، ص24).
- التحليل اللغوي
- التركيب اللغوي
- الإنتاج اللغوي
- التدقيق الآلي (السباعي، 2022، ص23).
- المعالجة الآلية (السباعي، 2022، ص19).

- المعاجم الآلية (السباعي، 2022، ص20).
- الترجمة الآلية (بشار، 2020، ص9).
- نمذجة اللغة للتعرف التلقائي على الكلام

3.4. اللسانيات الحاسوبية واللغة العربية:

إن البحث في مجال حوسبة اللغات عموماً واللغة العربية خصوصاً، قد اجتاز أشواطاً عديدة معبدة بالصعوبات والتحديات، ومع ظهور حواسيب الجيل السادس وتطور الدراسات في مجال هندسة اللغة، تجاوز طموح الإنسان المعاصر حوسبة اللغات فقط إلى البحث عن برامج محاكية للدماغ البشري وللسلوك الإنساني في مختلف جوانبه النفسية والحركية والتواصلية...، وذلك بصنع آلات أو روبوتات مفكرة لها القدرة والكفاية على فهم اللغة وإنتاجها في سياقات جديدة (برينيس وبشار، 2022).

وهنا يجب ذكر مجموعة من الباحثين اللسانيين الذين أثروا الساحة العلمية ببحوثهم في مجال اللسانيات الحاسوبية لقراءة جهودهم في المجال مثل علي حلمي موسى، نبيل علي، عبد ذياب عجيلي، محمد مراياتي، نهاد الموسى، عبد الرحمن الحاج صالح (شتوح وبلحداد، 2021).

يعد كتاب "اللغة العربية والحاسوب ل "نبيل علي" أول مؤلف يتناول موضوع اللسانيات الحاسوبية مطبقة على أنظمة اللغة العربية. للتأكيد على علاقة اللغة بالحاسوب، وإخضاع الحاسوب للغة لا العكس، من خلال الانطلاق من اللغة، ومن ثمة التفصيل في فروع اللغة العربية، وبربطها بالمعالجة الآلية. فهذا الكتاب يمثل القاعدة الأساسية للبحث العربي في مجال اللسانيات الحاسوبية (شتوح وبلحداد، 2021).

وترى الباحثة بأنه بالرغم من أن جهود علماء اللغة العرب في مجال اللسانيات الحاسوبية عديدة ومتنوعة لكنها أنها لم ترتقِ إلى المستوى المطلوب لمواكبة العصر الحالي ولم تتبناها المؤسسات العلمية والتربوية كمجالاً بحثياً يستحق تخصيص الموارد والاستثمار، فيُلاحظ أن بحوث اللغويات الحاسوبية العربية لا تزال في مرحلة النشأة مقارنة بنظيراتها الغربية التي شهدت تقدما فائقا في مجالاتها التطبيقية؛ لذلك ينبغي على المؤسسات العلمية والتربوية أن تضع استراتيجيات لتطوير المحتوى العربي بإدخال الرقمنة، وإنتاج محتوى هادف يسمو بالمستوى الثقافي للغة العربية، والحفاظ على مكانتها ومواكبتها للتطور التقني، حتى تبقى اللغة العربية صامدة ومنسجمة مع احتياجات أبنائها أمام التقدم التكنولوجي في كل أشكاله.

3.5. الذكاء الاصطناعي:

هو مصطلح صاغه البروفيسور بجامعة ستانفورد جون مكارثي في عام (1955)، على أنه علم وهندسة صنع آلات ذكية. يشير إلى تطوير أنظمة الكمبيوتر القادرة على أداء المهام التي تتطلب عادةً الذكاء البشري، مثل التعلم والاستدلال وحل المشكلات والإدراك وفهم اللغة. تستخدم أنظمة الذكاء الاصطناعي الخوارزميات والبيانات لاتخاذ القرارات وتصنيف الأشياء وتوليد الأفكار، غالباً دون برمجتها صراحة (McCarthy، 1959).

وبعد الاطلاع على التعريف الوارد في المرجع السابق وغيرها من المصادر العلمية، تعرف الباحثة الذكاء الاصطناعي بأنه مجال بحثي في علوم الحاسوب يطور ويدرس الأساليب والبرامج التي تمكن الآلات من إدراك بيئتها واستخدام التعلم والذكاء لاتخاذ إجراءات تزيد من فرصها في تحقيق أهداف محددة.

3.6. معالجة اللغة الطبيعية:

معالجة اللغة الطبيعية "هي مجموعة من التقنيات الحسابية ذات الدوافع النظرية لتحليل وتمثيل النصوص الطبيعية على مستوى واحد أو أكثر من مستويات التحليل اللغوي بغرض تحقيق معالجة لغوية شبيهة بالمعالجة البشرية لمجموعة من المهام أو التطبيقات" (Liddy, 2001, p1).

وعموماً، ترى الباحثة أن معالجة اللغة الطبيعية هي مجال يدرس اللغة البشرية كبنية رسمية، ويمكن اعتبارها حقل فرعي من الذكاء الاصطناعي وحقل متعدد التخصصات يعتمد على علوم الكمبيوتر واللغويات والإحصاء والاحتمالات، والهدف النهائي هو مساعدة الآلات في فهم اللغة ومعالجتها وإنتاجها بكفاءة على قدم المساواة مع البشر. كما استفادت معالجة اللغة الطبيعية من التطورات الأخيرة في الذكاء الاصطناعي والتعمق في بعض مناهجه كالتعلم الآلي (ML) والتعلم العميق(DL) ، فشهدت السنوات الخمس الماضية وحدها وفرة من تطبيقات معالجة اللغة الطبيعية المفيدة المتاحة لعامة الناس.

3.7. لمحة عن تطور معالجة اللغة الطبيعية:

لقد خضعت معالجة اللغة الطبيعية لتحولات كبيرة على مر السنين، حيث تطورت من نهج قائم على القواعد إلى نهج أكثر اعتماداً على البيانات ومبني على الشبكات العصبية. في هذه النظرة العامة، تتوضح المراحل الثلاث الرئيسية لتطور معالجة اللغة الطبيعية.

- معالجة اللغة الطبيعية الرمزية (1950–1980):

تميزت المرحلة الأولى من معالجة اللغة الطبيعية بنهج رمزي، اعتمد على قواعد وتمثيلات مشفرة يدوياً لمعالجة اللغة الطبيعية وفهمها. استند هذا النهج إلى فكرة أنه يمكن صياغة اللغة باستخدام مجموعة من القواعد والرموز. ركزت معالجة اللغة الطبيعية الرمزية على تطوير القواعد النحوية والمحللات والتمثيلات الدلالية لتحليل اللغة وتوليدها. نجح النهج الرمزي في تطوير أنظمة معالجة اللغة الطبيعية المبكرة، مثل الترجمة الآلية وأنظمة الإجابة على الأسئلة. ومع ذلك، فقد كانت لها حدود، بما في ذلك عدم القدرة على التعامل مع الغموض وعدم اليقين في اللغة، وصعوبة التوسع إلى مجموعات بيانات كبيرة، والقدرة المحدودة على التعلم من البيانات. وعلى الرغم من القيود، فقد أرست الأساس لأساليب البرمجة اللغوية العصبية المتطورة (Liddy, 2001, p5–7, 11–12).

- معالجة اللغة الطبيعية الإحصائية (1990–2000):

تميزت المرحلة الثانية من معالجة اللغة الطبيعية بتقديم الأساليب الإحصائية، والتي أحدثت ثورة في هذا المجال من خلال تمكين الآلات من التعلم من مجموعات البيانات الكبيرة. استخدمت معالجة اللغة الطبيعية الإحصائية خوارزميات التعلم الآلي لتحليل بيانات اللغة ونمذجتها، بدلاً من الاعتماد على قواعد مشفرة يدوياً.

أدت معالجة اللغة الطبيعية الإحصائية إلى تقدم كبير في نمذجة اللغة، ووسم أجزاء الكلام، والتعرف على الكيانات، واعتمدت على استراتيجيات تركز على البيانات لتعزيز فهم اللغة. ومع ذلك كانت معالجة اللغة الطبيعية الإحصائية لها حدودها الخاصة، بما في ذلك الاعتماد على كميات كبيرة من بيانات التدريب المصنفة، والقدرة المحدودة على التعامل مع الكلمات خارج المفردات، والفشل في التقاط العلاقات

الدلالية بين الكلمات. وكون هذا النهج نقيضاً للقواعد المحددة مسبقاً، فقد شجع على التعامل الفعال مع اللغة (Wachsmuth, 2023, p7–9).

- معالجة اللغة الطبيعية العصبية (منذ عام 2010 وحتى الآن):

تتميز المرحلة الثالثة من معالجة اللغة الطبيعية بظهور الشبكات العصبية، التي غيرت المجال من خلال تمكين الآلات من تعلم الأنماط المعقدة في بيانات اللغة. تستخدم معالجة اللغة الطبيعية العصبية خوارزميات التعلم العميق لتحليل بيانات اللغة ونمذجتها، غالباً مع الحد الأدنى من التدخل البشري.

أدت تقنيات معالجة اللغة الطبيعية العصبية إلى تقدم كبير في فهم اللغة وتوليدها والتعلم متعدد المهام، وتمكنت أيضًا من تطوير تضمين الكلمات، وآليات الانتباه، والتعلم بالتحويل (Zhang et al., 2016, p1).

ومما سبق وجدت الباحثة أن تطور معالجة اللغة الطبيعية تميز بتحولات كبيرة، من نهج رمزي إلى نهج إحصائي وأخيراً إلى نهج عصبي. وقد بنيت كل مرحلة على المرحلة السابقة، مما مكن الآلات من فهم أفضل للغة الطبيعية وتوليدها. ومع استمرار تطور معالجة اللغة الطبيعية، يمكن توقع رؤية المزيد من التطورات المثيرة في هذا المجال.

3.8. مهام معالجة اللغة الطبيعية الشائعة:

فيما يأتي قائمة ببعض المهام الأكثر بحثاً في معالجة اللغة الطبيعية. بعض هذه المهام لها تطبيقات مباشرة في العالم الحقيقي، في حين أن البعض الآخر يعمل بطريقة أكثر شيوعاً كمهام فرعية تُستخدم للمساعدة في حل مهام أكبر:

- **التحليل الصرفي:**

هو دراسة البنية الداخلية للكلمات وكيفية تكوينها من وحدات أصغر. يعد التحليل الصرفي مهم في معالجة اللغة الطبيعية لأنه يساعد في مهام تحديد الكلمات الجذرية والبادئات واللاحقات التي تشكل الكلمة ووضع علامات على أجزاء الكلام والتعرف على الكيانات المسماة ونمذجة اللغة، فضلاً عن تحليل العلاقات بين الكلمات ذات البنى الصرفية المتشابهة. من خلال تحليل البنية الصرفية لكلمة ما، يمكن لنموذج التعلم الآلي فهماً أفضلاً لمعناها وسياقها، وخاصة للغات ذات الصرف المعقد، مثل العربية (Liddy, 2001, p8).

- **التحليل النحوي:**

هو دراسة كيفية دمج الكلمات لتشكيل الجمل والهياكل النحوية الأخرى. يعد التحليل النحوي مهم في معالجة اللغة الطبيعية لأنه يساعد في مهام مثل تحليل الجملة وتحليل المشاعر وتحليل العلاقات بين الكلمات والترجمة الآلية وتحديد الأخطاء النحوية واقتراح التصحيحات. من خلال تحليل البنية النحوية للجملة، يمكن لنموذج التعلم الآلي فهماً أفضلاً للعلاقات بين الكلمات وتحديد الجملة الرئيسية والجمل الثانوية (Karoo; Jogi, 2023, p1).

- **الدلالات المعجمية:**

هي دراسة معنى الكلمات الفردية في السياق. تعد الدلالات المعجمية مهمة في المعالجة اللغوية العصبية لأنها تساعد في مهام مثل استنباط معنى الكلمة بما في ذلك المرادفات والمتضادات والاختصارات وتحليل المشاعر والإجابة على الأسئلة وتحديد غموض معنى الكلمة وحلها (Liddy, 2001, p8).

- **الدلالات العلائقية:**

هي دراسة العلاقات بين الكيانات والمفاهيم في الجمل الفردية. تعد الدلالات العلائقية مهمة في معالجة اللغة الطبيعية لأنها تساعد في مهام استخراج المعلومات والإجابة على الأسئلة وتلخيص النص وتحديد العلاقات بين الكيانات بما في ذلك أدوارها وسماتها وعلاقاتها واستنتاج علاقات جديدة (Turney, 2005, p1–2).

- **تحليل الخطاب:**

هو دراسة العلاقات بين الجمل وكيف تسهم في المعنى العام للنص. يعد تحليل الخطاب مهم في معالجة اللغة الطبيعية لأنه يساعد في مهام تلخيص النص وتحليل المشاعر وتحليل العلاقات بين الجمل بما في ذلك تماسكها وترابطها وبنيتها الخطابية والترجمة الآلية وتحديد نية المؤلف ونبرته. من خلال تحليل بنية الخطاب في النص، يمكن لنموذج التعلم الآلي فهماً أفضلاً للعلاقات بين الجمل وتحديد الموضوع الرئيسي والموضوعات الفرعية (Liddy, 2001, p10).

- **معالجة النصوص والكلام:**

تعد معالجة النصوص والكلام مجالاً أساسياً من مجالات معالجة اللغة الطبيعية، وتساعد في مهام معالجة وتحليل بيانات النصوص والكلام وفهمها وتلخيصها وتقسيم النصوص الأولية إلى أجزاء واستخراج جذورها وإزالة الكلمات الشائعة (أدوات الربط والتعريف وحروف العطف وحروف الجر والضمائر...) وعلامات الترقيم وتمثيل النص كمتجهات رقمية والتعرف على الكلام وتوليفه (Behera & Nayak, 2020, p2–3).

3.9. نماذج اللغة الكبيرة ونماذج اللغة الكبيرة المدربة مسبقاً:

نماذج اللغة الكبيرة هي فرع من فروع علوم الحاسوب والذكاء الاصطناعي الذي يهتم بالتفاعل بين الحاسوب واللغة البشرية. إنها دراسة النمذجة الرياضية والحسابية لمختلف جوانب اللغة وتطوير ترسانة من الأنظمة. تعد نماذج اللغة الكبيرة مجالاً للبحث والتطبيق يدرس كيف يمكن استخدام أجهزة الحاسوب لفهم ومعالجة نص أو كلام اللغة الطبيعية لأداء مهام مفيدة (Mohammad et al., 2023).

ويمكن النظر إلى نماذج اللغة الكبيرة بأنها خوارزميات تعلم عميقة تستخدم نماذج المحولات وتدرب باستخدام مجموعات بيانات ضخمة يمكنها تنفيذ مجموعة متنوعة من مهام معالجة اللغة الطبيعية. تحتوي نماذج اللغة الكبيرة على عدد كبير من المعلمات، والتي تشبه الذكريات التي يجمعها النموذج في أثناء تعلمه من التدريب.

تُشار عادةً إلى نماذج اللغة المدربة مسبقاً ذات أحجام المعلمات الأكبر وبيانات التدريب المكثفة باسم نماذج اللغة الكبيرة. يتجاوز حجم النموذج عادةً (6–10) مليارات معلمة. أصبح تدريب (LLMs) ضرورياً للغاية ويتطلب تدريب ونشر خبرة في التعامل مع البيانات واسعة النطاق وخبرة عملية كبيرة في التدريب. يؤكد هذا المطلب على الحاجة إلى أن يمتلك الباحثون الذين يطورون (LLMs) قدرات هندسية كبيرة في معالجة التحديات التي يواجهونها في أثناء عملية تطوير (LLM) (Liu, Y; et al., 2024).

وبعد الاطلاع على العديد من المراجع تعرف الباحثة نماذج اللغة الكبيرة المدربة مسبقاً بأنها نوع من نماذج الذكاء الاصطناعي التي دربت على مجموعة ضخمة من بيانات النصوص، صممت هذه النماذج لتعلم الأنماط والعلاقات في اللغة، ويمكن ضبطها بدقة لمهام محددة مثل ترجمة اللغة أو الإجابة على الأسئلة أو إنشاء

النصوص. يمكن أن تحتوي أكبر النماذج على حوالي 1 تريليون معلمة وتتطلب آلاف وحدات الرسوميات المتطورة للمعالجة وأسابيع أو أشهر من التدريب. وهذا يجعل التدريب المسبق مكلفاً للغاية وميسور التكلفة فقط لعدد قليل من الشركات والمنظمات في العالم التي تمتلك الموارد اللازمة.

3.10. أنواع نماذج اللغة الكبيرة:

يمكن تصنيف أنواع نماذج اللغات الكبيرة حسب عوامل مختلفة كالآتي:

3.10.1. حسب النطاق اللغوي:

- النماذج أحادية اللغة: دربت هذه النماذج على فهم وإنشاء نص بلغة واحدة، مثل النماذج المصممة للغة الفرنسية، مثل (FlauBERT) (Xu et al., 2024, p2).
- النماذج متعددة اللغات: تم تصميمها للتعامل مع لغات متعددة، وتدرب على مجموعات بيانات تمتد عبر لغات متنوعة لضمان قابلية التطبيق والعدالة عبر مختلف اللغات مثل عائلة نماذج (GPT) (Xu et al., 2024, p1).
- النماذج اللهجية: تعد نسخة متخصصة من نماذج اللغة تركز على الاختلافات اللغوية الإقليمية أو الاجتماعية. وهي أكثر تركيزاً من النماذج متعددة اللغات، حيث تتناول الخصائص اللغوية الفريدة للهجات (Mousi et al., 2024, p2–3) مثل (MARBERT) مع اللهجات العربية.
- النماذج التي تركز على لغات البرمجة: تتخصص هذه النماذج في لغات البرمجة، تسعى لتقديم مجموعة واسعة من مهام المطور، بما في ذلك إنشاء التعليمات البرمجية، وإنشاء الوثائق من التعليمات البرمجية، وإنشاء حالات

الاختبار، وإصلاح الأخطاء، وفهم مساحة العمل وحل الاستعلامات مثل Copilot (Agarwal et al., 2024, p1).

3.10.2. حسب التخصص في المهمة:

- النماذج العامة: صممت هذه النماذج، لتكون متعددة الاستخدامات وتؤدي مجموعة واسعة من المهام، من إنشاء النصوص إلى التلخيص والترجمة والمزيد، وغالباً ما تعتمد على الضبط الدقيق أو التحفيز لمهام محددة.
- النماذج المخصصة: مصممة خصيصاً لصناعات أو مجالات محددة أو مهام محددة، مثل (BioBERT) للنصوص الطبية الحيوية أو (FinBERT) للمستندات المالية تركز هذه النماذج على المعرفة المتخصصة وتضبط بدقة على مجموعات البيانات المتخصصة أو تطبيق مهمة محددة كالتلخيص أوالترجمة أو الإجابة على الأسئلة على سبيل المثال (PEGASUS) و(NMT) (Yang et al., 2020 & Lee et al., 2019).

3.10.3. حسب بنية النموذج:

تنقسم إلى العديد من الأنواع بناء على خوارزميات بناء النماذج، ستوضح بالتفصيل لاحقاً.

3.10.4. حسب نمط الإدخال:

- النماذج النصية فقط: تعالج هذه النماذج مدخلات النص الخالصة، المصممة حصرياً للمهام اللغوية (Wu et al., 2023, p1).
- النماذج متعددة الوسائط: قادرة على التعامل مع النص والأنماط الأخرى مثل الصور أو الصوت أو الفيديو (Wu et al., 2023, p1).

- نماذج تحويل الكلام إلى نص/نص إلى كلام: تتخصص هذه النماذج في معالجة إدخال/إخراج الصوت بالتزامن مع النص، مثل (Whisper) من (OpenAI) (OpenAI, 2022).

3.10.5. حسب الحجم والكفاءة:

- النماذج الكبيرة: نماذج كثيفة الموارد للغاية، وغالباً ما تحتوي على مليارات المعلمات، مثل (BERT)، تتفوق هذه النماذج في الأداء ولكنها تتطلب قوة حسابية كبيرة (Sanh et al., 2020, p1).
- النماذج متوسطة الحجم: توازن النماذج ذات المعلمات الأقل بين الأداء والكفاءة، مثل (DistilBERT)، مما يجعلها مناسبة للتطبيقات متوسطة المستوى (Sanh et al., 2020, p1).
- النماذج الصغيرة: صممت الإصدارات خفيفة الوزن مثل (TinyBERT) أو (MobileBERT) للأجهزة الطرفية والبيئات ذات الموارد المنخفضة، مما يضحي ببعض الدقة من أجل السرعة والكفاءة (Sun et al., 2020, p1).

3.10.6. حسب نهج التدريب:

تنقسم إلى العديد من الأنواع بناء على خوارزميات تدريب النماذج، ستوضح بالتفصيل لاحقاً.

3.10.7. حسب إمكانية الوصول:

- النماذج مفتوحة المصدر: تُعَد نماذج اللغة مفتوحة المصدر نماذج لغوية يكون الكود المصدري الخاص بها متاحاً للعامة ويمكن لأي شخص الوصول إليه واستخدامه وتعديله وتوزيعه بحرية. تسمح طبيعة المصدر المفتوح بقدر أكبر من الابتكار ومشاركة المعرفة وجهود التطوير الجماعية. حيث إن بعض العلماء يجادلوا أن الذكاء الاصطناعي مفتوح المصدر من شأنه أن يؤدي إلى

ابتكار أسرع، مشيرين إلى النجاح التاريخي لمبادرات مفتوحة المصدر أخرى مثل Linux، مما يدعو إلى التعاون العالمي وزيادة عدد الباحثين وتوليد الشفافية والثقة من خلال منح المزيد من الأشخاص القدرات على فحص وتدقيق النماذج والبحث عن التحيز وإخفاقات الموثوقية واكتشاف مشاكل عدم محاذاة القيمة (Potter et al, 2024).

- النماذج مغلقة المصدر: تعد نماذج لغة البرمجة ذات المصدر المغلق نماذج لغوية لا يتوفر كودها المصدري للعامة. تطور وتصان بوساطة منظمات أو شركات تحافظ عادةً على ملكية الكود الأساسي وإغلاقه أمام الجمهور. يرى البعض أن نماذج الذكاء الاصطناعي المغلق المصدر من شأنه أن يؤدي إلى ابتكار أسرع، نظراً لأن الشركات التي تحتفظ بحقوق الملكية لنماذجها تحافظ أيضاً على حافز أكبر للاستثمار الخاص. يمكن لهذه الشركات دعم فرق التطوير المخصصة التي لديها حافز تجاري لنشر التكنولوجيا بسرعة أكبر، ولا تزال قادرة على تحقيق التعاون من خلال تقاسم العمل فيما بينها، حيث يبررون لك بأنه يمكن أن يمنع من توحيد الأفكار والأساليب، على النقيض من المخاطر المحتملة للتقارب عندما يقوم الأفراد أو المنظمات بنسخ ولصق نماذج المصدر المفتوح (Potter et al, 2024).
- النماذج الهجينة: هي نوع من نماذج اللغات الكبيرة يجمع بين نقاط القوة في كل من (LLMs) مفتوحة المصدر ومغلقة المصدر. وهو يستخدم بنية مفتوحة المصدر كأساس، ولكنه يشتمل على مكونات خاصة، مثل بيانات التدريب أو خوارزميات التحسين من النماذج مغلق المصدر (Babenko, 2024).

3.11. تطبيقات نماذج اللغة الكبيرة:

وبعد اطلاع الباحثة على العديد من المراجع وجدت أن نماذج اللغة الكبيرة تستخدم في تطبيقات مختلفة عبر صناعات ومجالات متعددة. وفيما يأتي لخصت بعض الأمثلة على هذه التطبيقات:

– ترجمة اللغة: يمكن ضبط نماذج اللغة الكبيرة للترجمة الآلية، مما يتيح ترجمة دقيقة وفعالة للنص من لغة إلى أخرى.

– فهم اللغة: يمكن استخدام نماذج اللغة الكبيرة لتحليل وفهم اللغة الطبيعية، مما يتيح تطبيقات مثل التعرف على الكلام وتصنيف النص والتعرف على الكيانات المسماة.

– تعلم اللغة: يمكن استخدام نماذج اللغة الكبيرة لإنشاء أدوات تعلم لغة مخصصة، ويمكن استخدامها لإنشاء ألعاب قائمة على اللغة، مثل ألعاب تعلم اللغة أو ألعاب الكلمات.

– تلخيص النص: يمكن لنماذج اللغة الكبيرة تلخيص النصوص الطويلة إلى ملخصات موجزة وذات مغزى، مما يوفر الوقت والجهد للقراء.

– تصنيف النص: يمكن لنماذج اللغة الكبيرة تصنيف النص إلى فئات، مثل رسائل البريد الإلكتروني العشوائية وغير العشوائية، أو مراجعات المنتجات الإيجابية والسلبية.

– تحليل المشاعر: يمكن لنماذج اللغة الكبيرة تحليل النص لتحديد المشاعر أو النبرة العاطفية وراءه، وهو أمر مفيد في تطبيقات مثل تحليل تعليقات العملاء ومراقبة وسائل التواصل الاجتماعي لفهم المشاعر والآراء العامة حول مواضيع مختلفة.

– إنشاء المحتوى: يمكن لنماذج اللغة الكبيرة إنشاء محتوى عالي الجودة، مثل المقالات ومنشورات المدونات ومنشورات وسائل التواصل الاجتماعي، وأتمتة إنشاء

المحتوى لمختلف الصناعات، وإنشاء المحتوى الإبداعي مثل الشعر أو القصص أو حتى الكتب بأكملها، ويمكن استخدامها لتوصية المحتوى للمستخدمين بناءً على اهتماماتهم وتفضيلاتهم.

– روبوتات الدردشة والمساعدون الافتراضيون: يمكن لنماذج اللغة الكبيرة تشغيل روبوتات الدردشة والمساعدون الافتراضيون، واستخدامها لتشغيل أنظمة خدمة العملاء الآلية، مما يمكنهم من فهم استفسارات المستخدم والرد عليها بطريقة أكثر إنسانية ويقدم دعماً أكثر كفاءة وفعالية للعملاء.

– الإجابة على الأسئلة: يمكن ضبط نماذج اللغة الكبيرة للإجابة على الأسئلة، مما يتيح لها الاستجابة بدقة بناءً على السياق والمحتوى.

هذه مجرد أمثلة قليلة للتطبيقات العديدة لنماذج اللغة الكبيرة. يتطور المجال بسرعة، وتكتشف تطبيقات جديدة وتطويرها باستمرار.

3.12. كيف تعمل نماذج اللغة الكبيرة:

تعمل نماذج اللغة الكبيرة من خلال سلسلة من الخطوات التي تمتد من معالجة البيانات والتدريب إلى الاستدلال والضبط الدقيق. فيما يأتي شرح خطوة بخطوة لكيفية عمل نماذج اللغة الكبيرة:

1. تحديد الهدف: يجب أن تكون هناك حالة استخدام نماذج اللغة الكبيرة، وسيؤثر الهدف في مصادر البيانات التي سيسحب منها. يمكن أن يتطور الهدف وحالة استخدام نماذج اللغة الكبيرة لتشمل عناصر جديدة مع تدريب النماذج وضبطها.

2. جمع البيانات: تدرب نماذج اللغة الكبيرة على مجموعات بيانات ضخمة، غالباً ما تتضمن الكتب ومواقع الويب والمقالات وغيرها من مجموعات النصوص الكبيرة،

لالتقاط مجموعة واسعة من هياكل اللغة والموضوعات والمعرفة (Liu, Y; et al., 2024, p6).

3. الترميز والمعالجة المسبقة: تنظف بيانات النص وتقسم إلى وحدات صغيرة (رموز) مثل الكلمات، مما يساعد النموذج في التعامل مع المفردات والسياق بكفاءة والمعرفة (Liu, Y; et al., 2024, p8).

4. تدريب النموذج: تُعد نماذج التدريب جانباً بالغ الأهمية يؤثر في فاعليتها. يمكن لعوامل مثل كمية ونوعية بيانات التدريب وهندسة النموذج وخوارزميات التعلم المستخدمة أن تؤثر بدرجة أكبر في أداء وتعميم نماذج (LLM) (Management Solutions, 2024).

5. إعداد بنية النموذج: حالياً تبنى جميع نماذج (LLM) على بنية المحول، مما يسمح لهذه النماذج بالتوسع إلى مليارات عدة أو حتى تريليونات معلمة. تنقسم أنواع البنيات إلى ثلاث فئات (Encoder-only) و(Encoderdecoder) و(Decoder-only). لم تعد بنية (Encoder-only) مستخدمة في أحدث نماذج (LLM) (Liu, Y; et al., 2024, p9). تتضمن بنية المحول طبقات الاهتمام الذاتي وطبقات التغذية الأمامية والتي تكون مرتبة في مجموعات. يحتاج نماذج (LLM) إلى موارد حسابية مثل جهاز كمبيوتر قوي أو خادم قائم على السحابة للتعامل مع عمليات التدريب. غالباً ما تحد متطلبات الموارد هذه من قدرة العديد من المنظمات على تطوير النماذج الخاص بهم.

6. الضبط الدقيق (اختياري): هي عملية ضبط النموذج لجعله مناسب لتطبيقات محددة، وتحسن عمل نماذج التعلم اللغوي (Elastic, 2024) عن طريق التدريب بدقة على مجموعات بيانات خاصة بالمهمة واستراتيجيات عدة لمواءمة أوثق للمخرجات مع

التوقعات البشرية للجودة والملاءمة، على سبيل المثال الترجمة والتلخيص والإجابة على الأسئلة.

7. الاستدلال: يشير الاستدلال في نماذج اللغة الكبيرة إلى عملية استخدام نموذج مدرب لتوليد نص أو التنبؤ بالرمز التالي في تسلسل، بناءً على الأنماط التي تعلمها النموذج (Chatgptguide, 2024) مع الأخذ في الاعتبار مدخلات أو سياق معين. الاستدلال هو المرحلة الأخيرة التي يستخدم فيها النموذج (Hazelcast, 2024)، هذه العملية أساسية لعمل (LLMs) وهي ما يسمح لها بالتفاعل مع المستخدمين بطريقة ذات مغزى.

8. التقييم: لقد أصبح تقييم هذه النماذج أكثر تعقيداً، مما يتطلب النظر في المشاكل والمخاطر المحتملة من جميع الجوانب (Liu, Y; et al., 2024, p16). منذ شعبية (ChatGPT)، نشرت العديد من الدراسات ذات الصلة، وهو أمر مفيد لتطوير نماذج التعلم العميق واسعة النطاق وهذه الخطوة هي موضع هذا البحث.

3.13. مراحل تدريب نماذج اللغة الكبيرة:

يعد التدريب المسبق مرحلة أساسية في تطوير نماذج اللغة الكبيرة، وعلى الرغم من أن هذه العملية تتطلب الكثير من الحساب وتكلف الكثير (Zhou et al., 2024, p1)، إلا أنها تمكن النموذج من التكيف مع مجموعة واسعة من المهام.

بعد اطلاع الباحثة على المراجع الحديثة وماورد فيها وجدت أن عملية تدريب نموذج لغوي مسبقاً هي عملية هامة جداً تتم على مجموعة كبيرة من البيانات لضبطه بدقة لمهمة محددة. ارتأت من خلال البحث لوضع خطوات مراحل تدريب النموذج بأنها تتضمن جمع البيانات ومعالجتها، تحديد بنية النموذج، وتحديد تقنيات التدريب المناسبة، وخطوة الضبط الدقيق والمحاذاة.

3.13.1. جمع البيانات ومعالجتها:

البيانات هي الأساس الذي تُبنى عليه نماذج اللغة الكبيرة، وتؤثر جودتها وتنوعها وتمثيلها مباشرة في أداء وانحياز النموذج الناتج، إن معالجة التحديات المتعلقة بالملكية الفكرية أمر ضروري لتطوير نماذج قوية وغير متحيزة ودقيقة (Management Solutions, 2024).

وجدت الباحثة أن مجموعة البيانات التدريبية هي مجموعة بيانات نصية تستخدم لتدريب نماذج اللغة الكبيرة. تتضمن هذه المجموعة مصادر متنوعة مثل الكتب ومواقع الويب والأوراق الأكاديمية ووسائل التواصل الاجتماعي والمستندات القانونية وبيانات المحادثة.

3.13.1.1. صفات مجموعات البيانات:

يجب أن تتصف مجموعة البيانات هذه بصفات عديدة لتساعد في تحقيق أفضل أداء للنموذج:

- تنوع المحتوى: تساعد مجموعة البيانات الغنية والمتنوعة نماذج اللغة الكبيرة في التعميم، وتوفر فهماً دقيقاً عبر العديد من المجالات (Liu Y. et al., 2024, p18)، كما تعكس مجموعة واسعة من الأصوات والآراء والخبرات. تستفيد نماذج اللغة من المحتوى الذي يغطي مواضيع وأساليب ونبرات متنوعة، بما في ذلك الكتابة الرسمية والمحادثات غير الرسمية والمواد الفنية أو الخاصة بالمجال.
- تعدد اللغات: لضمان قدرة النموذج على التعامل مع لغات وسياقات ثقافية متعددة (Liu Y. et al., 2024, p5)، غالباً ما تتضمن المجموعة نصاً

بلغات ولهجات مختلفة، بالإضافة إلى مواد متنوعة ثقافياً. يساعد هذا في منع التحيز تجاه أي لغة واحدة أو منظور ثقافي.

- الأهمية الزمنية: تصبح مجموعة البيانات قديمة بسرعة، قد تكون التحديثات المنتظمة ضرورية (Yu et al., 2024, p5) للحفاظ على الأهمية الزمنية في النماذج الأحدث.
- الملكية الفكرية وحقوق النشر: غالباً ما تتضمن النصوص الضخمة المستخدمة في تدريب نماذج اللغة الكبيرة محتوى خاصاً أو محمياً بحقوق الطبع والنشر، مما يثير مخاوف أخلاقية وقانونية نظراً لأنه لا يجوز استخدامها أو توزيعها دون إذن. يحتاج مزودو نماذج اللغة الكبيرة إلى التأكد من أن لديهم الحقوق أو التراخيص لاستخدام هذه النصوص لتجنب انتهاك القانون (Yu et al., 2024, p6).
- الدقة والموثوقية: يجب أن تكون بيانات التدريب عالية الجودة دقيقة وخالية من الأخطاء وجديرة بالثقة (Yu et al., 2024, p5)، مما يقلل من المعلومات المضللة والتحيزات. يمكن أن تؤدي البيانات منخفضة الجودة إلى توليد النماذج لاستجابات غير صحيحة أو مضللة.
- المراجعة البشرية: من المعروف أن أداء نماذج اللغة الكبيرة يمكن تحسينه عن طريق إدخال ردود الفعل البشرية في المراحل المختلفة من تطويرها (Li et al., 2024, p3). ونظراً لعدم تمكن هذه النماذج مهما وصلت من تطور وإنجاز على جميع المستويات ولا سيما المستوى اللغوي، لكنها تبقى غير قادرة على إنجاز أداء كالإنسان؛ لذلك لا بد من تواجد إشراف بشري مثل التعليق والتحقق، يؤدي إلى تحسين جودة البيانات بدرجة أكبر ويضمن دقة أعلى وتحيزاً أقل.

3.13.1.2. سياسات معالجة البيانات:

من خلال عمل الباحثة وجدت أن عملية تحضير مجموعة البيانات تتضمن مجموعة من سياسات معالجة البيانات وتصفيتها قبل بناء النموذج. غالباً ما تحتوي بيانات التدريب على ضوضاء، مثل التكرار أو الجمل غير المكتملة أو المعلومات غير ذات الصلة أو اللغة غير المرغوب فيها. تتضمن المعالجة المسبقة سياسات لإزالة الضوضاء وتعزيز الاتساق وتحسين كفاءة التدريب وجودة المخرجات.

تنظيف البيانات:

إن معالجة البيانات مسبقاً لنماذج اللغة الكبيرة (LLMs) أمر بالغ الأهمية لضمان جودة وأهمية البيانات التي تتعلم منها هذه النماذج. وفيما يأتي نظرة تفصيلية على أهمها:

- تجنب اللغة غير القياسية: الابتعاد عن العامية واللهجات قدر الإمكان عند اختيار بيانات التدريب.
- إزالة العناصر غير الضرورية: قد يتضمن ذلك الأحرف والرموز الخاصة والعناصر غير النصية مثل الصور أو الجداول التي لا تفيد في بيانات اللغة (Liu, Y; et al., 2024, p8).
- الأخطاء الشائعة والإملائية والمطبعية: تساعد خوارزميات تصحيح الإملاء والقواعد الأساسية في تحديد هذه الأخطاء والمساعدة في تحييد آثارها حيث يمكن أن تؤدي إلى إعادة النموذج لاستجابات خاطئة، مما يؤدي إلى تنبؤات أقل دقة.

- إزالة التكرارات: يمكن للبيانات المكررة أن تحرف عملية تعليم النموذج من خلال الإفراط في تمثيل عبارات أو هياكل معينة. يضمن تحديد التكرارات وإزالتها مجموعة بيانات متوازنة (Liu, Y; et al., 2024, p8).
- هيكلة البيانات الصحيحة: التأكد من أن مجموعة بيانات التدريب متناسقة وصحيحة.

الترميز (Tokenization):

يشير الترميز إلى عملية تقسيم النص إلى وحدات أصغر تسمى "الرموز"، وهي الوحدات التي يعالجها برنامج (LLM) في أثناء عملية التدريب واستنتاج الاستجابة. يمكن أن تكون هذه الرموز كلمات أو أجزاء من الكلمات أو أحرفاً (Management Solutions, 2024).

توجد خوارزميات ترميز عدة تختلف في طريقة تقسيم النص يُذكر منها:

- (BytePairEncoding): هي طريقة لتقسيم الكلمات الفرعية إلى رموز تستخدم في نماذج التعلم العميق لأنها تسمح للنموذج بالتعرف على وحدات أصغر على مستوى أجزاء الكلمة والتي تدمج في كلمات أكبر (Sennrich et al., 2015).
- (SentencePieceEncoding): هي طريقة تقسيم الكلام إلى أجزاء تولد رموزاً فرعية للكلمات. تعامل الجملة بأكملها كسلسلة واحدة دون الحاجة إلى مسافات صريحة بين الكلمات مما يجعلها مستقلة عن اللغة ومثالية للغات التي لا تُستخدم فيها المسافات (مثل الصينية أو اليابانية) (Kudo & Richardson, 2018).

- التقسيم إلى وحدات على مستوى الكلمة (Word-level tokenization): يقسم النص إلى كلمات فردية، وهو ما قد يعمل جيداً للغات ذات حدود الكلمات الواضحة (مثل اللغة الإنجليزية) ولكنه قد يواجه صعوبة في اللغات المعقدة (مثل اللغة العربية) (Schuster & Nakajima, 2012).
- التقسيم إلى وحدات على مستوى الأحرف (Character-level tokenization): يمكن أن تكون التجزئة على مستوى الأحرف مفيدة في حالات مثل التعامل مع الأخطاء المطبعية أو الرموز الفريدة (Fleshman & Van Durme, 2023).

ولاحظت الباحثة من خلال بحثها ومراجعتها أنه يجب الأخذ بعين الاعتبار بعض الأمور في أثناء عملية الترميز:

- حجم المفردات: يعد اختيار حجم المفردات المناسب أمراً بالغ الأهمية، لأنه يوازن بين تعقيد النموذج والكفاءة الحسابية.
- الكفاءة: تعد سرعة التجزئة مهمة لمجموعات البيانات الكبيرة، ويمكن استخدام أدوات التجزئة المتوازية لتسريع العملية.
- القدرة على التعامل مع الرموز الخاصة: غالباً ما تقدم رموز مثل علامات الترقيم، أو المحارف الخاصة لمساعدة النموذج في فهم أجزاء مختلفة من المدخلات.

التضمين (Embedding):

هو تمثيل رقمي مكتسب لكلمات أو عبارات أو حتى جمل كاملة، والفكرة الأساسية وراء التضمينات هي رسم خريطة لرموز منفصلة في مساحة توضع فيها الكلمات أو العبارات المتشابهة دلالياً بالقرب من بعضها البعض (Management

(Solutions, 2024, p24. تكمن أهميتها بأنها تسمح للخوارزميات بالعمل مع بيانات نصية بتنسيق يمكن للنماذج معالجته بفعالية.

التكميم (Quanitification):

يشير التكميم عادةً إلى تقنية تستخدم لتقليل متطلبات الحوسبة والذاكرة للنموذج من خلال تمثيل أوزانه (المعلمات الرقمية التي تعلمها النموذج) بعدد أقل من Bits (Management Solutions, 2024, p26). وهذا يسمح بمعالجة أسرع واستخدام أقل للذاكرة والقدرة على نشر نماذج كبيرة على أجهزة ذات موارد محدودة، مثل الأجهزة المحمولة أو الأجهزة الطرفية دون التأثير بدرجة أكبر في أداء النموذج والتضحية بالدقة بدرجة أكبر.

التطبيع (Normalization):

هو تقنية تستخدم لتثبيت التدريب، وتحسين تقارب النموذج، وضمان بقاء التنشيطات والتدرجات ضمن نطاق يمكن التحكم فيه. إنها تساعد النماذج الكبيرة في التعلم بطريقة أكثر فعالية، وتؤدي إلى أداء أفضل وأوقات تدريب أسرع (Ba et al., 2016, p1).

لخصت الباحثة أبرز تقنيات التطبيع كالآتي:

- تصغير الأحرف: يُحول كل النص إلى أحرف صغيرة لمنع النموذج من التمييز بناءً على الحالة.
- التقسيم إلى أجزاء وتبسيط الكلمات: يختصر الكلمات إلى أشكالها الأساسية لتوحيد أشكال مختلفة من نفس الكلمة.
- التوحيد: تحويل التمثيلات المختلفة للأحرف إلى نموذج قياسي، مما يمنع النموذج من تفسيرها ككيانات منفصلة.

- توحيد الاختصارات والانقباضات: يضمن الاتساق، مما يساعد النموذج في فهم أفضل لأنماط اللغة.

تسهم كل من جوانب إعداد بيانات التدريب في تطوير نماذج اللغة الكبيرة الأخلاقية والمتوافقة مع القانون والأداء العالي والتنوع، وتلعب دوراً أساسياً في تحديد قدراتها الشاملة وموثوقيتها.

3.13.2. تحديد بنية النموذج (خوارزميات بنى نماذج اللغة الكبيرة):

يمكن تجميع خوارزميات بنى نماذج اللغة الكبيرة على نطاق واسع في الفئات الآتية:

1. النماذج القائمة على الشبكة العصبية المتكررة (RNNs Recurrent Neural Networks):

إن الشبكات العصبية المتكررة من بين أقدم الهياكل المعمارية المستخدمة للبيانات المتسلسلة، وخاصة في نمذجة اللغة. تعالج هذه النماذج البيانات في تسلسل، مما يسمح لها بتذكر المعلومات من الرموز السابقة، والتي يمكن أن تكون مفيدة لمهام مثل إنشاء النصوص والترجمة (Amazon, 2023).

- (Long Short-Term Memory) (LSTMs):

هي طريقة جديدة وفعالة تعتمد على التدرج، هذا لا يسبب ضرراً، يمكن للذاكرة الطويلة قصيرة المدى أن تتعلم سد الفجوات الزمنية الدنيا التي تتجاوز عدد خطوات زمنية منفصلة من خلال فرض تدفق ثابت للخطأ من خلال دوامات الأخطاء الثابتة داخل وحدات خاصة. كما تحل أيضاً المهام المعقدة الاصطناعية ذات التأخير الزمني

الطويل والتي لم تحل أبداً بوساطة خوارزميات الشبكة المتكررة السابقة لها (1997 ,Hochreiter & Schmidhuber).

تعتمد نوعاً من الشبكات العصبية المتكررة وهي مصممة للتعامل مع البيانات القائمة على التسلسل، ولا تزال تستخدم في بعض أنظمة الذاكرة طويلة المدى، وخاصة تلك التي تتطلب إدارة ذاكرة أكثر وضوحاً. فإن (LSTM) تؤدي إلى المزيد من التشغيلات الناجحة، وتتعلم بسرعة أكبر.

- GRUs (Gated Recurrent Units):

هي نوع من بنية الشبكة العصبية المتكررة تشبه (LSTM)، وهي مصممة للتعامل مع البيانات المتسلسلة والاعتماديات الطويلة المدى. قدمت GRUs كبديل أبسط وأكثر كفاءة حسابياً ل (LSTM)، وخاصة لنمذجة التسلسلات في مهام مثل معالجة اللغة الطبيعية والتعرف على الكلام وتحليل البيانات الأخرى المعتمدة على الوقت. اقترح هذا المنهج لجعل كل وحدة متكررة تلتقط التبعيات لمقاييس زمنية مختلفة. وعلى غرار وحدة (LSTM)، تحتوي وحدة (GRU) على وحدات بوابات تعدل تدفق المعلومات داخل الوحدة، ولكن دون وجود خلايا ذاكرة منفصلة (Chung et al., 2014).

2. النماذج القائمة على شبكات التغذية الأمامية (FFNs Feedforward Networks):

تعمل الشبكات العصبية ذات التغذية الأمامية بدون حلقات تغذية مرتدة (Sazli, 2006)، وتستخدم عموماً بنيات مثل الطبقات المتصلة بالكامل أو المحولات. في (LLM) غالباً ما تستخدم (FFNs) داخل بنيات المحولات لمعالجة الرموز في وقت

واحد، بدلاً من التتابع، مما يسمح بمزيد من التوازي ومعالجة أفضل للتبعيات طويلة المدى (Pires et al., 2023).

3. المحولات (Transformers):

هي بنية نموذجية تتجنب التكرار وتعتمد بدلاً من ذلك بالكامل على آلية الانتباه لرسم التبعيات بين المدخلات والمخرجات. يسمح المحول بقدر أكبر من التوازي ويمكنه الوصول إلى حالة جديدة من الفن في جودة الترجمة بعد تدريبه لمدة اثنتي عشرة ساعة فقط على ثماني وحدات معالجة رسومية. إن اعتماد المحول على آلية الانتباه الذاتي يسمح لها بمعالجة المعلومات وربطها من أي جزء من التسلسل بأي جزء آخر دون الحاجة إلى معالجة متسلسلة، كما هو الحال في البنيات المتكررة مثل (LSTM) و(GRUs) (Vaswani et al., 2017).

تخلصت الباحثة إلى أن المحولات هي بنية أساسية للشبكات العصبية لنماذج اللغة الكبيرة (LLMs) ويُنظر إليها على نطاق واسع باعتبارها المحفز للعديد من التطورات في معالجة اللغة الطبيعية في السنوات الأخيرة لأنها عالية الكفاءة وقابلة للتوازي وقادرة على التقاط العلاقات المعقدة في النص، مما يجعلها مثالية لنماذج اللغة الكبيرة الحديثة.

4. النماذج القائمة على التعلم التعزيزي (RL Reinforcement Learning):

يطبق التعلم التعزيزي عادةً نظام الجوائز في سياق ضبط النماذج لتحقيق أهداف محددة أو تحسين الاستجابات بناءً على الملاحظات. يساعد ذلك في محاذاة مخرجات النموذج مع المعايير الخاصة بالهدف (Havrilla et al., 2024). لا تعمل تقنيات التعلم التعزيزي على إنشاء بنية النموذج، بل تعمل بدلاً من ذلك على تكييفها من خلال حلقات الملاحظات.

- (RLHF Reinforcement Learning from Human Feedback):

يستخدم في تدريب نماذج اللغة الكبيرة لتحسين أدائها في المهام الذاتية أو الدقيقة من خلال دمج التفضيلات البشرية في عملية التعلم الخاصة بالنموذج. فهو يجمع بين مبادئ التعلم التعزيزي والملاحظات البشرية لضبط النماذج، وخاصة للمهام التي يكون من الصعب الحصول على الإشراف المباشر في الممارسة العملية (Du et al., 2024)، يتضمن ذلك "مكافأة" المخرجات التي تتوافق مع الاستجابات المرغوبة و"معاقبة" غير المرغوب فيها.

- RL for Chatbots and Interactive Systems:

يتعلم فيها النموذج كيفية تحسين استجاباته بناءً على ردود الفعل التفاعلية، بهدف تعزيز رضا المستخدم أو ملاءمة الاستجابة أو نجاح المهمة. هذا النهج الذي يستخدم غالباً في ضبط نماذج اللغة الكبيرة، يمكن النموذج من التعلم من أخطائه وتعديل سلوكه ديناميكياً استجابةً لردود الفعل البشرية (Linkedin, 2024).

3.13.3. تدريب النموذج (خوارزميات تدريب نماذج اللغة الكبيرة):

هناك العديد من خوارزميات تدريب نماذج اللغة، يُذكر منها:

1- نمذجة تسلسل إلى تسلسل: حيث يدرب النموذج لتوليد تسلسلات نصية بناءً على تسلسلات إدخال أخرى (Management Solutions, 2024).

2- نمذجة اللغة المقنعة (Masked Language Modeling): والتي تتكون من إخفاء بعض الكلمات عشوائياً في نص الإدخال وتدريب النموذج للتنبؤ بهذه الكلمات المقنعة بناءً على السياق المحيط. تسمح هذه التقنية بالتعلم ثنائي الاتجاه وفهم أفضل للسياق. ويوجد استراتيجية أكبر تدعى النطاق المقنع هذه التقنية مشابهة لـ MLM،

ولكن بدلاً من إخفاء الرموز الفردية، تخفى نطاقات النص بالكامل (Management Solutions, 2024).

3- النمذجة اللغوية الانحدارية الذاتية أو النمذجة أحادية الاتجاه (Autoregressive Language Modeling or Unidirectional Modeling): والتي تتألف من تدريب النموذج للتنبؤ بالكلمة التالية (Next Sentence Prediction) أو مقطع النص المعطى للسياق السابق. تتيح هذه المهمة للنموذج تعلم الاحتمالات الشرطية للغة وإنشاء نص جيد (Management Solutions, 2024).

4- النمذجة غير الانحدارية: حيث لا يحصل على الاستجابة بتسلسل كلمة بكلمة، بل تحول وتنقح ككل (Management Solutions, 2024).

5- التعلم الخاضع للإشراف (Supervised Learning): يستخدم بيانات مسماة، حيث يكون لكل مدخل مخرجات مرتبطة (Label)، مما يسمح للنموذج بالتعلم من الأمثلة. يقلل النموذج من الاختلاف بين تنبؤاته والتسميات الفعلية. غالباً ما يستخدم هذا النوع من التدريب في مرحلة الضبط الدقيق لتخصيص (LLM) لمهام محددة مثل تصنيف النص أو تحليل المشاعر أو التعرف على الكيانات المسماة (Named Entity Recognition) (Alpaydin, 2014).

6- التعلم غير الخاضع للإشراف (Unsupervised Learning): يستخدم التعلم غير الخاضع للإشراف عندما لا توجد تسميات. يتعلم النموذج الأنماط والهياكل والعلاقات في البيانات بناءً على النص المدخل فقط (Hinton & Sejnowski, 1999).

7- التعلم الذاتي الإشرافي (Self-Supervised Learning): يولد تسميات زائفة من خلال التلاعب بالبيانات نفسها، مثل إخفاء الكلمات أو خلط ترتيب الجمل، ثم يتعلم النموذج التنبؤ بهذه التسميات المولدة (Jaiswal et al., 2021).

8- التعلم التعزيزي (RL): الفكرة وراء سياسة التدريب هذه استخدام تدرجات السياسات أو التعلم لتحسين مخرجات النموذج.

9- التعلم التعزيزي من ردود الفعل البشرية (Reinforcement Learning from Human Feedback): تستخدم ردود الفعل البشرية لتشكيل سلوك النموذج. ويتعلم من خلال تلقي المكافآت أو العقوبات بناءً على أفعاله، حيث يهدف إلى تعظيم المكافآت التراكمية لتحسين أدائه في مهمة محددة.

10- التعلم بالتحويل (Transfer Learning): يستفيد من المعرفة المستمدة من نموذج مدرب مسبقاً على مهمة واحدة لتحسين الأداء في مهمة مختلفة ومرتبطة غالباً. يعمل النموذج المدرب مسبقاً كأساس ويضبط بدرجة أكبر على مجموعة بيانات أصغر ذات صلة بالمهمة الجديدة، مثل نماذج معالجة اللغة الطبيعية القانونية أو الطبية (Pan & Yang, 2010).

11- التعلم بالمناهج (Curriculum Learning): يبدأ التعلم بمهام أسهل ويتقدم تدريجياً إلى مهام أكثر تعقيداً، محاكياً الطريقة التي يتعلم بها البشر. يساعد هذا النموذج في استقرار التدريب وتجنب التعثر في الأمثلة الصعبة. هذا النوع من التدريب مفيد خاصة لتحسين التقارب والأداء، وخاصة في النماذج المدربة من الصفر أو على مجموعات بيانات كبيرة جداً (Bengio et al., 2009).

12- التعلم المستمر (Continual Learning): يتضمن تحديث النموذج تدريجياً ببيانات جديدة عند توافرها دون إعادة تدريب النموذج بالكامل أو فقدان المعرفة من

البيانات التي تعلمها مسبقاً. هذا أمر صعب بسبب النسيان الكارثي (Catastrophic Forgetting)، حيث ينسى النموذج المعرفة السابقة عند تعرضه لبيانات جديدة (De Lange et al., 2021).

13- التعلم المقارن (Contrastive Learning): تتضمن هذه التقنية تدريب النموذج لزيادة التشابه بين النص المدخل وإصداراته الموسعة، مع تقليل التشابه مع النصوص الأخرى غير ذات الصلة (Zhang et al, 2024, p8).

14- التعلم متعدد المهام (Multi-Task Learning): يُدرب النموذج على مهام متعددة ذات صلة في وقت واحد. من خلال تعلم التمثيلات المشتركة عبر المهام، يصبح النموذج أكثر قوة وقابلية للتعميم (Sener & Koltun, 2019).

15- التعلم النشط (Active Learning): يحدد التعلم النشط الأمثلة الأكثر إفادة للتصنيف، مما يزيد من تحسين النموذج بأقل قدر من البيانات المصنفة. غالباً ما يستخدم التعلم النشط في السيناريوهات حيث يكون الحصول على البيانات المصنفة مكلفاً أو يستغرق وقتًا طويلاً، مما يحسن أداء النموذج مع عدد أقل من التعليقات التوضيحية (Settles, 2009).

16- التعلم الفوقي (Meta Learning): هو أسلوب لتدريب النموذج على تعلم كيفية التعلم، من خلال التدريب على مجموعة من المهام ثم ضبطه بدقة على مهمة جديدة (Finn, 2017).

17- التعلم الجماعي (Ensemble Learning): يجمع بين نماذج متعددة أو نفس النموذج المدرب بتكوينات مختلفة لإجراء تنبؤات. تقلل التنبؤات المجمعة من تحيزات النموذج الفردي وتحسن الدقة الإجمالية. يستخدم في بعض المهام التي تتطلب موثوقية أو دقة عالية، كما هو الحال في معالجة اللغة الطبيعية الطبية (Zhou, 2012).

18- نقاط التفتيش والتوقف المبكر (Checkpointing and Early Stopping): تحفظ حالة النموذج على فترات في أثناء عملية التدريب، ويوقف التوقف المبكر التدريب إذا لم يتحسن أداء التحقق بمرور الوقت. يساعد ذلك في منع الإفراط في التجهيز وتوفير طريقة للعودة إلى إصدار سابق من النموذج إذا انخفض الأداء بسبب الإفراط في التدريب (Prechelt, Lutz, 2002).

19- التقطير وضغط النموذج (Distillation and Model Compression): هو تقنية ضغط النموذج حيث ينقل نموذج مدرب أكبر (Teacher) المعرفة إلى نموذج أصغر (Student). يتعلم النموذج الأصغر تقريب مخرجات المعلم، مع الاحتفاظ بمعظم أدائه. يستخدم التقطير لإنشاء إصدارات فعالة من نماذج التعلم الأساسية التي يمكن تشغيلها على أجهزة أقل قوة (مثل الأجهزة المحمولة) أو الاستجابة بسرعة أكبر في بيئات التشغيل (Production Environments) (Hinton et al., 2015).

20- زيادة البيانات (Data Augmentation): تولد زيادة البيانات، بيانات إضافية عن طريق تعديل عينات البيانات الموجودة. في معالجة اللغة الطبيعية، قد يشمل هذا إعادة الصياغة أو حذف الكلمات العشوائية أو الترجمة العكسية. يساعد ذلك في تحسين التعميم والحد من الإفراط في التجهيز وتعزيز الأداء، وخاصة في البيئات ذات الموارد المنخفضة حيث تكون البيانات المصنفة محدودة (Zou & Wei, 2019).

21- Few-Shot and Zero-Shot Learning: يستخدم التعلم بطريقة FewDhot بيانات محدودة للغاية لمهمة جديدة، ولا يستخدم التعلم بطريقة Zero-Shot أي بيانات خاصة بالمهمة، ويعتمد بالكامل على المعرفة المدربة مسبقاً للنموذج (Brown et al., 2020).

22- Prompt Engineering: تُستخدم الأوامر المصممة بعناية لاستنباط النتائج المرغوبة دون تدريب إضافي. إنها تستفيد من المعرفة الموجودة مسبقاً للنموذج للتكيف مع مهام أو سياقات مختلفة.

23- تقنيات التنظيم (Regularization Techniques): يعد الإفراط في التجهيز تحدياً رئيسياً في بناء النموذج، يوجد عدد من تقنيات التنظيم لتقليل ذلك وتحسين التعلم الانتقالي، منها تقنية التسرب وتنظيم L2 نحو أولوية خارج النطاق، بالإضافة إلى ذلك tuneout، وهي تقنية مستوحاة من التسرب (Barone et al, 2017, p1).

24- التدريب التنافسي: تتضمن هذه التقنية تدريب النموذج بزيادة بيانات تدريب الشبكة العصبية ليكون قوياً ضد الهجمات التنافسية، مثل اضطرابات الإدخال أو تسميم البيانات. لقد ثبت باستمرار أن التدريب التنافسي، يعزز من متانة النموذج ضد الهجمات المختلفة. ويُحسِّن التعميم وزيادة الأداء في مهام معالجة اللغة الطبيعية (Xhonneux et al., 2024, p2-3).

25- الشبكات التنافسية التوليدية (Generative Adversarial Networks): تدريب النموذج باستخدام شبكة مولدة وشبكة تمييز، لتوليد نص لا يمكن تمييزه عن النص الحقيقي، مع تدريب جهاز تمييز أيضاً لاكتشاف النص الناتج (Goodfellow et al., 2014).

26- الشبكات العصبية البايزية (Bayesian Neural Networks): توفر إطاراً صارماً لتحليل وتدريب الشبكات العصبية التي تدرك عدم اليقين، وعموماً، لدعم تطوير خوارزميات التعلم. وهو يعتمد على فكرتين بسيطتين، إن أول هذه الأفكار هو أن الاحتمالية هي مقياس للاعتقاد بحدوث أحداث، وليس الحد الأقصى لتكرار حدوثها

عندما يتجه عدد العينات نحو اللانهاية. أما الفكرة الثانية فهي أن الاعتقادات السابقة تؤثر على الاعتقادات اللاحقة (Jospin et al., 2022, p1).

27- التشفير التلقائي المتغير (Variational Autoencoders): يمكن استخدام الاستدلال الخلفي التقريبي المتعلم لمجموعة من المهام مثل التعرف وإزالة الضوضاء والتمثيل والتصور. عندما يستخدم شبكة عصبية لنموذج التعرف، نصل إلى المشفر التلقائي المتغير والذي يستخدم كتقنية تدريب (LLMs) (Kingma, Welling, 2022).

غالباً ما يجمع بين هذه التقنيات وضبطها لتحقيق نتائج متطورة في مهام معالجة اللغة الطبيعية المختلفة، حيث لكل منها نقاط قوة ونقاط ضعف خاصة بها، ويعتمد اختيار التقنية على المهمة المرجوة والتطبيق المراد.

3.13.4. الضبط الدقيق والمحاذاة Fine-tuning:

نماذج اللغة المضبوطة بدقة هي نماذج لغوية كبيرة مستمدة من بنيتها الأساسية. وهي مخصصة لحالات استخدام أو مجالات محددة، وبالتالي تصبح أفضل في أداء المهام الأكثر تخصصاً. فإن النماذج المضبوطة بدقة يمكنها أداء مهام محددة بدرجة أكبر من النماذج الأساسية، فإن قوتها الأكبر هي أنها أخف وزناً، وأسهل تدريباً عموماً. (Chockalingam et al., 2023).

وتعرفه الباحثة بأنه شكل من أشكال التعلم بالنقل، حيث يستفيد النموذج من المعرفة العامة التي اكتسبها في أثناء عملية التدريب المسبق على مجموعة ضخمة من النصوص ويصقل هذه المعرفة للتفوق في مهمة أضيق وأكثر تركيزاً.

3.13.4.1. كيفية عمل الضبط الدقيق:

تتبع عملية الضبط الدقيق مراحل جمع البيانات والمعالجة المسبقة اختيار مجموعات البيانات الخاصة بالهدف وتنفيذ طرائق المعالجة المسبقة الفعالة. تختار طريقة الضبط الدقيق وفق النموذج المدرب المختار، وتضبط المعلمات الفائقة وفقاً لذلك (Jeong, 2024, p11).

وجدت الباحثة من خلال الاطلاع على الدراسات السابقة والمراجع المذكورة في هذا البحث أن عملية الضبط الدقيق تبدأ بنموذج لغة كبير مدرب على مجموعة بيانات متنوعة وكبيرة الحجم، حيث يستطيع النموذج فهم اللغة العامة، دون أي مهمة محددة في الاعتبار. ثم يجب استخدام مجموعة بيانات أصغر ومُسمَّاة ذات صلة بالمهمة المحددة التي نريد أن يؤديها النموذج والتي يحدث باستخدامها.

وتستخدم عملية الضبط الدقيق عادةً خوارزميات التعلم الخاضع للإشراف، ويمكن اختيار تحديث النموذج بالكامل أو أجزاء محددة فقط. يُضبط النموذج لتقليل الخطأ بين تنبؤاته والنواتج الحقيقية عن طريق مجموعة البيانات الخاصة بالمهمة. بعد الانتهاء من عملية الضبط الدقيق، يُقيم النموذج على مجموعة التحقق ونشره للاستخدام في العالم الحقيقي.

3.13.4.2. أنواع الضبط الدقيق:

- الضبط الدقيق تحت الإشراف (Supervised Fine-Tuning): الطريقة الأكثر شيوعاً، حيث يُضبط النموذج باستخدام بيانات مُسمَّاة لمهمة محددة. وبينما تكون هذه الطريقة فعّالة، فإنها تتطلب بيانات مُصنَّفة كبيرة، والتي قد تكون مكلفة وتستغرق وقتاً طويلاً للحصول عليها (Parthasarathy et al., 2024, p9).

- الضبط الدقيق من غير إشراف (Unsupervised Fine-Tuning): لا تتطلب هذه الطريقة بيانات مُصنَّفة. وبدلاً من ذلك، يتعرض النموذج لمجموعة كبيرة من النصوص غير المُصنَّفة من المجال المستهدف، مما يُحسِّن فهمه للغة. يُعد هذا النهج مفيداً للمجالات الجديدة مثل المجالات القانونية أو الطبية، ولكنه أقل دقة بالنسبة للمهام المحددة مثل التصنيف أو التلخيص (Parthasarathy et al., 2024, p9).
- الضبط الدقيق للتعليمات (Instruction Fine-Tuning): هو عملية تدريب النموذج على مجموعة بيانات تتكون من أزواج (تعليمات، مخرجات). تساعد هذه العملية في سد الفجوة بين هدف التنبؤ بالكلمة التالية لـ (LLM) وهدف المستخدم المتمثل في جعله يتبع التعليمات البشرية (Guo, 2024, p1).
- الضبط الدقيق في عدد قليل من المحاولات (Few-Shot Fine-Tuning): يضبط باستخدام بيانات مُسمَّاة محدودة للغاية. يستخدم عدد قليل من الأمثلة لمهمة تصنيف مع تحقيق أداء جيد بسبب التدريب المسبق القوي للنموذج. ينصب التركيز في هذا النوع من التعلم على استخدام مجموعات البيانات الصغيرة لتدريب النموذج. فإن جمع كمية كبيرة من البيانات الموثوقة لا يزال مكلفًا ويستغرق وقتًا طويلاً، أو حتى مستحيلًا. في هذا الصدد، ثبت أن التعلم من خلال اللقطات القليلة يتمتع بمزايا من حيث معالجة عينات البيانات الصغيرة غير الطبيعية في مساحة التطبيق الضخمة (Duan et al., 2021, p1).
- الضبط الدقيق القائم على المحول (Adapter-Based Fine-Tuning): تعمل هذه الطريقة عن طريق إضافة وحدات محولات خفيفة الوزن إلى نموذج لغة مدرب مسبقاً وتحديث معلمات وحدات المحول فقط عند التعلم في مهمة

لاحقة (He et al., 2021, p1) مع الحفاظ على التدريب السابق دون تغيير إلى حد كبير. هذا أكثر كفاءة في استخدام الذاكرة وخاصة للمواقف التي تحتاج فيها إلى ضبط نفس النموذج لمهام متعددة.

3.14. تطبيقات استخدام نماذج اللغة الكبيرة:

تكمن أهمية نماذج اللغة الكبيرة بالأثر الكبير لها في مختلف المجالات، ويُذكر من أهمها:

- التعليم: تعمل نماذج اللغة الكبيرة على إحداث ثورة في التعليم من خلال تمكين دعم الطلاب والمعلمين وتطوير المحتوى التعليمي. تساعد في تعزيز تجارب التعلم الشخصية من خلال تحليل أنماط التعلم والأداء والتفضيلات الخاصة بهم وتسهيل إمكانية الوصول. أما بالنسبة للمعلمين، يمكن لنماذج اللغة الكبيرة المساعدة في إنشاء خطط الدروس وتصحيح الواجبات وإنشاء محتوى تعليمي متنوع وشامل، مما يوفر بدرجة أكبر المزيد من الوقت للتدريس والتفاعل مع الطلاب (Naveed et al., 2024).
- الآثار الاجتماعية والأخلاقية: تفرض تحديات مثل خصوصية البيانات والتحيز وإساءة الاستخدام المحتملة. ويتطلب معالجة هذه المخاوف ممارسات شفافة للبيانات وحوكمة أخلاقية للذكاء الاصطناعي وأطر تنظيمية. ويعد ضمان العدالة والقدرة على التفسير أمراً ضرورياً للتخفيف من تأثيرها الاجتماعي (Vavekanand, Kumar, 2024).
- العلم: يمكن لنماذج اللغة الكبيرة تسريع عملية البحث من خلال تحليل وتلخيص الأدبيات العلمية بسرعة وطريقة مفهومة وسهلة الوصول. بالإضافة إلى ذلك، يمكنها مساعدة العلماء في قدرتهم على معالجة مجموعات البيانات واسعة النطاق مما يسمح لهم بالكشف عن أنماط قد لا تكون واضحة على

الفور للباحثين البشر. هذا لا يوفر الوقت فحسب، بل يحسن أيضاً وضوح الاتصال العلمي، مما يمكن الفرق متعددة التخصصات من العمل معاً بطريقة أكثر فعالية (Naveed et al., 2024).

تعمل نماذج اللغة الكبيرة على إعادة تشكيل الصناعات وإعادة تعريف التعاون بين الإنسان والذكاء الاصطناعي. وفي حين أن إمكاناتها هائلة، فإن معالجة التحديات المرتبطة بها ستكون مفتاحاً لتسخير فوائدها بمسؤولية واستدامة.

3.15. أهمية وفوائد نماذج اللغة الكبيرة:

لقد أصبح تطبيق نماذج اللغة الكبيرة على مجموعة متنوعة من المهام شائعة في كل من مجتمعات البحث والصناعات المرتبطة بالذكاء الاصطناعي، مع اكتشاف العديد من الاستخدامات الناشئة واستكشافها يومياً. تُلحظ تطبيقات نماذج اللغة الكبيرة في الطب والتعليم والعلوم والرياضيات والقانون والتمويل والروبوتات والترميز. في حين أن كل من هذه المجالات تشكل تحديات مختلفة، فإن نماذج اللغة الكبيرة تفتح فرصاً لتقديم مساهمات كبيرة في هذه المجالات (Naveed et al., 2024) يُذكر من فوائدها التي أسهمت في هذا الانتشار:

- الكفاءة: تعمل نماذج اللغة الكبيرة على تبسيط سير العمل من خلال أتمتة المهام المتكررة أو التي تستغرق وقتاً طويلاً، مما يتيح عمليات أسرع وأكثر كفاءة (Deng & Lin, 2022, p82). من خلال التعامل مع هذه المهام كل على حدة، تسمح نماذج اللغة الكبيرة للمحترفين بالتركيز على المساعي الاستراتيجية أو الإبداعية، مما يعزز الإنتاجية بدرجة أكبر، وتقلل أوقات الاستجابة وتعزز حل المشكلات من خلال توفير حلول فورية ودقيقة.

- التعميم: إن إحدى نقاط القوة الرئيسية هي قدرة هذه النماذج على تعميم المعرفة عبر مجموعة واسعة من المهام، من خلال التدريب على مجموعات بيانات متنوعة، دون الحاجة إلى تدريب محدد للمهمة. يمكن للنموذج التبديل بسلاسة بين المهام والتكيف مع حالات الاستخدام المختلفة (Reizinger et al., 2024).
- تحليل البيانات: هنا تتفوق النماذج اللغة الكبيرة؛ فهي قادرة على استخراج وتحليل المعلومات ذات المغزى من مجموعة البيانات (Cheung, 2024, p8). تتمتع برامج التعلم الآلي بفعالية لا تصدق في معالجة وتحليل كميات كبيرة من البيانات غير المنظمة واتخاذ القرارات بناء عليها، وتفيد في مختلف الصناعات.
- تعدد الوسائط: إن قدرة النماذج على استخدام مدخلات ومخرجات متعددة الوسائط تعمل على إثراء التطبيقات، ومن خلال ذلك تقدم رؤى أعمق وتمكن من تفاعلات أكثر ثراءً (Khalid et al.,2024)؛ حيث لا تقتصر تفاعلاتها على النص فقط، بل إنها تدمج أيضاً أنماط بيانات أخرى مثل الصور والصوت والفيديو.
- إمكانية الوصول: توفر نماذج اللغة الكبيرة وصولاً أوسع وأسهل إلى المعلومات والخدمات من خلال توفير واجهات بديهية وسهلة الاستخدام. يمكن للمستخدمين التفاعل مع الأنظمة باستخدام اللغة الطبيعية (Wang et al., 2024, p1) بدلاً من الأوامر المعقدة. كما أنها تعمل على تشغيل تقنيات مساعدة لذوي الاحتياجات الخاصة والمتحدثين بالعديد من اللغات.
- تحسين تجربة المستخدمين: تعمل نماذج اللغة الكبيرة على تحسين قطاعات خدمة الزبائن عن طريق تقديم تفاعلات أفضل وذات كفاءة عملية (Meduri,

2024, p1). فهي تعمل على تشغيل برامج المحادثة الآلية والمساعدين الافتراضيين الذين يقدمون دعم العملاء على مدار الساعة طوال أيام الأسبوع، مما يضمن حل الاستفسارات وتحسين الرضا ويعزز الشعور بالمشاركة والولاء تعزيزاً كبيراً من خلال استجابات مخصصة وواعية بالسياق في الوقت المناسب ويجعل التفاعلات أكثر طبيعية ومتعة للمستخدمين.

- قابلية التوسع: إن تدريب نماذج اللغة الكبيرة هو إجراء مكثف حسابياً يتطلب موارد كبيرة من الأجهزة والطاقة. ومن الضروري الوصول إلى مجموعات الحوسبة الفائقة أو الأجهزة المتخصصة من أجل تدريب النماذج الكبيرة (Raiaan et al., 2024, p29). مما وجه الباحثين لجعل أنظمة التعلم الآلي قابلة للتوسع؛ لتكون قادرة على التعامل مع أحمال العمل المتزايدة دون المساس بالأداء أو الحاجة إلى إصلاحات معمارية كبيرة. كما يعمل دعمها للغات متعددة على توسيع نطاقها، مما يمكن الشركات من خدمة شريحة أكبر من المستخدمين.
- خفض التكاليف: تتمثل إحدى الفوائد الرئيسية لنماذج اللغة الكبيرة في قدرتها على خفض التكاليف للشركات. نظراً لأن نماذج اللغة الكبيرة مدربة مسبقاً على مجموعات بيانات واسعة النطاق، فلا تحتاج المؤسسات إلى تدريب النماذج من الصفر، مما يوفر الموارد الحسابية ووقت التطوير والحاجة إلى فرق دعم كبيرة أو قوى عاملة يدوية وتسهم في التخلص من الحاجة إلى الاستثمارات الأولية في البنية التحتية الباهظة الثمن، مما يؤدي إلى خفض التكاليف (Aryan et al.,2023, p5).
- التخصيص: عملية تخصيص نماذج اللغة الكبيرة للأغراض العامة وفقاً لبيانات سياقية محددة، مع تعزيزها بهذه المعرفة، وتحسينها وفقاً للأهداف المرجوة،

وتنظيمها من خلال القيود المحددة للمجال (Ling et al., 2024, p2). تعد نماذج اللغة الكبيرة قابلة للتكيف بدرجة كبيرة ويمكن تخصيصها لمهام أو صناعات أو احتياجات تنظيمية محددة من خلال الضبط الدقيق أو الهندسة السريعة.

3.16. التحديات والاعتبارات التي تواجهها نماذج اللغة الكبيرة:

لقد حققت نماذج اللغة الكبيرة تقدماً كبيراً في الآونة الأخيرة. ومع توسيع نطاق هذه النماذج والتعلم المستمر والتعامل مع مهام أكثر تعقيداً، فإنها تجلب مجموعة من التحديات، يُذكر أهمها:

- التكلفة الحسابية: تتطلب نماذج اللغة الكبيرة موارد حسابية واسعة النطاق، مما يزيد من تكاليف الإنتاج ويثير المخاوف البيئية بسبب استهلاك الطاقة الكبير في أثناء عملية التدريب على نطاق واسع (Naveed et al., 2024).
- التأثيرات المجتمعية السلبية: تثير نماذج اللغة الكبيرة المخاوف نتيجة لحقيقة مفادها أنها قادرة على توليد محتوى مقنع للغاية ولكنه قد يكون كاذباً أو مضللاً. ويمكن استخدام هذا المحتوى لنشر معلومات مضللة، أو التلاعب بالآراء، أو إنشاء مواد احتيالية، الأمر الذي يخلف عواقب وخيمة على الرأي العام والديمقراطية والتلاحم الاجتماعي (Jiao et al., 2024, p9). ويمكن إساءة استخدامها لتوليد هجمات تصيد مقنعة، والتي يمكن أن تقوض ثقة الجمهور في مصادر المعلومات (Fred, 2023) وتؤدي إلى عواقب وخيمة لحياة الأفراد المستهدفين بهذه الهجمات.
- التحيز وعدم الإنصاف: يمكن أن تعمل نماذج اللغة الكبيرة على تأكيد التحيزات الموجودة في البيانات المستخدمة للتدريب. وقد يؤدي هذا إلى مخرجات متحيزة أو تمييزية تعزز التحيزات المجتمعية القائمة (Fred, 2023).

- مخاوف الخصوصية والأمان: أثار تطور نماذج اللغة الكبيرة الجدل حول أمن البيانات والخصوصية فيها، وإدراكاً متزايداً للحاجة الملحة لحماية المعلومات الحساسة. يحاول الباحثون إيجاد تقنيات مبتكرة للتخفيف من مخاطر الخصوصية المتأصلة في نماذج اللغة الكبيرة، والتي غالباً ما تعالج كميات هائلة من البيانات، بما في ذلك معلومات حساسة محتملة وقد تكون خاصة بمستخدمين (Jiao et al., 2024, p13).
- الإفراط في التجهيز: على الرغم من أن نماذج (LLM) تمتلك قدرات تعليمية كبيرة، إلا أنها معرضة للإفراط في التجهيز لأنماط صاخبة وغريبة في بيانات التدريب المكثفة الخاصة بها، وبالتالي قد يتسبب هذا في توليدها لاستجابات غير منطقية (Naveed et al., 2024).
- الهلوسة: تظهر برامج التعلم العميق "الهلوسة"، حيث تولد استجابات على الرغم من أنها تبدو معقولة، إلا أنها غير صحيحة أو لا تتوافق مع المعلومات المقدمة. يمكن تصنيف الهلوسة إلى ثلاث فئات (الهلوسة المتضاربة مع المدخلات، الهلوسة المتضاربة مع السياق، الهلوسة المتضاربة مع الحقائق) (Naveed et al., 2024).
- المعرفة القديمة: قد تحتوي المعلومات الواقعية التي تستخدم في أثناء عملية التدريب المسبق على أخطاء أو تصبح قديمة بمرور الوقت. ومع ذلك، فإن إعادة تدريب النموذج باستخدام بيانات ما قبل التدريب المحدثة أمر مكلف، ومحاولة إلغاء الحقائق القديمة وتعلم حقائق جديدة في أثناء عملية الضبط الدقيق ليست بالأمر السهل (Kaddour et al., 2023).
- العمالة: في حين أن نماذج اللغة الكبيرة لديها القدرة على خلق فرص عمل جديدة، إلا أنها تشكل أيضاً تهديداً كبيراً لأنواع معينة من التوظيف، وخاصة

تلك التي تنطوي على مهام روتينية ومتكررة مثل خدمة العملاء وإدخال البيانات وإنشاء المحتوى. كما يؤدي المزيد من الاعتماد على عمل الذكاء الاصطناعي إلى عدم المساواة في القوى العاملة، حيث قد يتمتع أولئك الذين لديهم المهارات والموارد للعمل مع النماذج بميزة على أولئك الذين لا يملكون.

- النسيان الكارثي: يفرض التعلم المستمر تحديات خاصة على الشبكات العصبية الاصطناعية بسبب ميل المعرفة بالمهام التي تم تعلمها سابقاً إلى الضياع فجأة مع دمج المعلومات ذات الصلة بالمهمة الحالية. تحدث هذه الظاهرة التي يطلق عليها النسيان الكارثي على وجه التحديد عندما تدرب الشبكة بتسلسل على مهام متعددة (Kirkpatrick et al., 2017).
- المعالجة في الوقت الفعلي: غالباً ما تحتوي نماذج اللغة الكبيرة على مئات الطبقات وملايين المعلمات، مما يعيق المعالجة في الوقت الفعلي بسبب المتطلبات الحسابية العالية وتخزين الوزن المحدود على منصات الأجهزة، تواجه تكلفة تنفيذ كبيرة بسبب العدد الكبير من طبقات النموذج، مما يؤدي إلى زمن انتقال استدلال مرتفع (Naveed et al., 2024).

3.17. التطورات المستقبلية في نماذج اللغة الكبيرة:

إن مستقبل نماذج اللغة الكبيرة مثير ويتطور بسرعة. وفيما يأتي بعض الاتجاهات والتطورات المحتملة التي تخلصت إليها الباحثة من خلال اطلاعها على المراجع العلمية المختلفة خلال إجراء هذا البحث والتي ارتأت أنها قد تشكل مستقبل نماذج اللغة الكبيرة:

- ستصبح نماذج اللغة الكبيرة معتمدة على نطاق واسع في مختلف الصناعات، بما في ذلك خدمة العملاء وإنشاء المحتوى وترجمة اللغات.

- سيكون هناك حاجة ملحة إلى المزيد من البحث لمعالجة التحديات الأخلاقية التي تفرضها نماذج اللغة الكبيرة وتقييم تأثيرها في العالم الحقيقي. ويشمل ذلك التحقيق في استراتيجيات التخفيف من التحيز في هذه النماذج، وتطوير آليات قوية لمنع إساءة الاستخدام، واستكشاف طرائق لضمان الشفافية والمساءلة (Bender et al., 2021).
- ستستمر نماذج اللغة الكبيرة في التحسن من حيث الدقة والطلاقة والتماسك، مما يجعلها أكثر فعالية في مجموعة متنوعة من التطبيقات.
- ستصبح نماذج اللغة الكبيرة أكثر تخصصاً، مع تدريب النماذج على مجالات محددة، مثل الطب أو القانون أو التمويل.
- ستكون نماذج اللغة الكبيرة قادرة على معالجة البيانات متعددة الوسائط وتوليدها، بما في ذلك النصوص والصور والصوت.
- ستصبح نماذج اللغة الكبيرة مستقلة، وقادرة على التعلم والتكيف من تلقاء نفسها دون تدخل بشري.

3.18. أشهر نماذج اللغة الكبيرة:

شكلت النماذج اللغوية التي طورت في السنوات الأخيرة أهم المحطات في تطور الذكاء الاصطناعي ووضعه في الاستخدام على نطاق واسع. قامت الباحثة بذكر أهمها بعد اطلاعها على المراجع المذكورة في هذا البحث ومواقع شركات البرمجة والذكاء الاصطناعي العالمية مثل OpenAI، Google، Microsoft، Anthropic، Nividia، DeepMind، Meta، Mistral AI:

1. (BERT 2018) (Bidirectional Encoder Representations from Transformers): نموذج لغوي مدرب مسبقاً طورته شركة Google وقد أحدث ثورة في مجال معالجة اللغة الطبيعية من خلال تقديم نهج جديد لنمذجة

اللغة، حيث يستفيد من بنية المحول ثنائي الاتجاه ليدرب النموذج على مهمة نمذجة لغة مقنعة، والتنبؤ بالكلمات المفقودة في الجملة. وهذا يسمح له بالتقاط العلاقات السياقية بين الكلمات وتحقيق نتائج متطورة في مهام معالجة اللغة (Devlin et al., 2018).

2. (XLNet 2019): نموذج لغوي مدرب مسبقاً طورته شركة Google وجامعة (Carnegie Mellon) وهو يعتمد على نجاح (BERT). لقد استخدم هدف تدريب جديد، يسمى نمذجة اللغة بالتبديل لالتقاط السياق ثنائي الاتجاه دون قيود نمذجة اللغة المقنعة، والذي يسمح للنموذج بالتنبؤ بالكلمة التالية في الجملة مع إعطاء تبديل عشوائي لتسلسل الإدخال، جاء هذا التحسن على حساب تعقيد حسابي أعلى (Yang et al., 2019).
3. (RoBERTa 2019) (Robustly Optimized BERT): أحد أشكال (BERT) الذي طورته شركة (Facebook AI). وهو يقدم العديد من التحسينات على بنية (BERT)، بما في ذلك هدف تدريب مختلف هو التدريب على مجموعات بيانات أكبر وإزالة مهمة التنبؤ بالجملة التالية واستخدام المزيد من التكرارات ووقت تدريب أطول. أظهر أداءً أعلى في معايير مختلفة مقارنةً بـ (BERT)، مما يوضح تأثير استراتيجيات التدريب المحسنة (Liu Y et al., 2019).
4. (T5 2019) (Text-to-Text Transfer Transformer): نموذج لغة مدرب مسبقاً طورته شركة Google وهو يؤطر جميع مهام معالجة اللغة الطبيعية كمهمة تحويل نص إلى نص، حيث يكون كل من المدخلات والمخرجات عبارة عن تسلسلات نصية. أظهر تنوعاً وأداءً استثنائياً عبر

مجموعة واسعة من تطبيقات معالجة اللغة الطبيعية، على الرغم من أنها تتطلب موارد حسابية كبيرة (Raffel et al., 2019).

5. (BART 2019) (Bidirectional and Auto-Regressive Transformers): نموذج لغة مدرب مسبقاً طورته شركة (Facebook AI) إنه يجمع بين نقاط القوة في (BERT) ونماذج تسلسل إلى تسلسل. لقد استخدمت مشفراً تلقائياً لإزالة الضوضاء وإعادة بناء النص الفاسد، مما يسمح له بالأداء الجيد في كل من مهام فهم اللغة الطبيعية والتوليد والتلخيص والترجمة (Lewis et al., 2019).

6. (OpenAI GPT Series 2020-Now) (Generative Pre-trained Transformer): نموذج لغوي واسع النطاق طورته (OpenAI) يولد نصاً متماسكاً وطبيعياً. بدأت سلسلة (GPT) ب (GPT-1) في عام (2018) تلاها (GPT-2) في عام (2019) و(GPT-3) في عام (2020) و(GPT-4) في عام (2023). وقد حققت هذه النماذج نمواً هائلاً في الحجم والقدرة، حيث تضم (GPT-3) (175) مليار معلمة. وهي تتفوق في مهام المحادثة وتطبيقات المجالات المتعددة. وفي حين أحدثت قدراتها ثورة في الذكاء الاصطناعي، ظهرت مخاوف بشأن التحيزات والمعلومات المضللة والاستخدام الأخلاقي جنباً إلى جنب مع نشرها.

7. (Megatron- Turing NLG 2021): نموذج لغوي واسع النطاق قائم على المحول طورته Microsoft وNVIDIA. حُسن لكفاءة التدريب على مجموعات وحدة معالجة الرسوميات، ويحقق نتائج متطورة في العديد من معايير معالجة اللغة الطبيعية، بما في ذلك ترجمة اللغة وتصنيف النص والتلخيص وترجمة اللغة.

8. (Gropher 2021): نموذج لغوي واسع النطاق قائم على المحول طورته شركة (DeepMind). ركز هذا النموذج على توسيع نطاق نماذج (LLM) مع التخفيف من التحديات مثل عدم دقة الحقائق والقضايا الأخلاقية. لقد سلط أداؤها الضوء على قيمة نماذج التوسع لتحسين التفكير وتمثيل المعرفة. لقد أرسى الأساس لتطوير الذكاء الاصطناعي المسؤول على نطاق كبير.
9. (Claude 2021): نموذج لغوي واسع النطاق قائم على المحول طورته شركة (Anthropic) مع التركيز على السلامة والمحاذاة والاعتبارات الأخلاقية في نمذج اللغة. دُرب (Claude) على توليد استجابات مفيدة وغير سامة، وأعطى الأولوية لإنتاج مخرجات آمنة وقابلة للتفسير لمهام المحادثة والحوار، على الرغم من أنه أخطأ في بعض الأحيان.
10. (LaMDA 2022) (Language Model for Dialogue Applications): نموذج لغوي واسع النطاق قائم على المحول طورته شركة Google مصمم خصيصاً لتطبيقات الحوار والمحادثة. صمم LaMDA للحفاظ على السياق وإنتاج استجابات جذابة تشبه الإنسان، وأعطى الأولوية لفهم نية المستخدم وتوليد إجابات ذات صلة وعقلانية.
11. (OPT 2022) (Open Pre-trained Transformer): نموذج لغوي واسع النطاق قائم على المحول طورته شركة (Meta AI) كبديل مفتوح المصدر لـ (LLM) للنماذج المغلقة المصدر. أعطى الأولوية للشفافية وإمكانية الوصول للباحثين مع الحفاظ على الأداء التنافسي في مهام معالجة اللغة الطبيعية. أشار تطويره إلى دفع نحو إضفاء الطابع الديمقراطي على الوصول إلى تكنولوجيا الذكاء الاصطناعي.

12. (PaLM 2022) (Pathways Language Model): نموذج لغوي واسع النطاق قائم على المحول طورته شركة Google مصمم للاستدلال والتعلم من لقطات قليلة وتطبيقات متعددة اللغات. مع 540 مليار معلمة، أظهر (PaLM) أداءً متطوراً في مهام الاستدلال المعقدة وأظهر قدرات استثنائية عند التعامل مع تعددية اللغات.

13. (LLaMA 2023): نموذج لغوي واسع النطاق قائم على المحول طورته شركة (Meta AI) هو نموذج أصغر نطاقاً ولكنه عالي الأداء لأغراض البحث. مع نماذج تتراوح من (7) مليارات إلى (65) مليار معلمة، قدمت (LLaMA) سهلة الوصول وفعالة ومُحسَّنة للاستخدام الأكاديمي.

14. (Gemini 2023): نموذج لغة كبير متقدم من (Google DeepMind) المصمم للتفوق في التطبيقات المتعددة الوسائط، حيث يجمع بين القدرات عبر النصوص والصور والصوت والفيديو والترميز. قُدم (Gemini) كخليفة لـ (PaLM 2)، ويمثل قفزة كبيرة في تكنولوجيا (LLM) من خلال تقديم تعدد الوسائط والأداء المتطور في التفكير وفهم اللغة والمهام الإبداعية.

15. (Mistral 2023): في أواخر عام (2023)، كشفت (Mistral AI) عن نماذجها اللغوية المعيارية والفعالة، بما في ذلك منها يتبع بنى الهندسة الكثيفة والمختلطة الخبيرة. أكدت (Mistral) على الحلول المرنة وعالية الأداء لمهام معالجة اللغة الطبيعية، والتي تلبي احتياجات الباحثين والمطورين الذين يسعون إلى نماذج الذكاء الاصطناعي القابلة للتطوير.

3.18.1. أمثلة عن نماذج اللغة الكبيرة الداعمة والمخصصة للغة العربية:

طُورت نماذج اللغة العربية الآتية لمعالجة التعقيدات والفروق الدقيقة للغة العربية. يُظهر التطور الزمني اهتماماً متزايداً بالشمول والحجم، مما يتيح مجموعة واسعة من التطبيقات في معالجة اللغة الطبيعية للغة العربية؛ قامت الباحثة بذكر أهمها بعد اطلاعها على المراجع المذكورة في هذا البحث ومواقع شركات البرمجة والذكاء الاصطناعي العالمية مثل Google، OpenAI:

1. (AraBERT 2020): طُور في عام (2020)، كان أول نموذج لغوي واسع النطاق يركز على اللغة العربية. بني على بنية (BERT) الخاصة بشركة (Google) ولكن دُرب حصرياً على النص العربي، بما في ذلك اللغة العربية الفصحى الحديثة واللهجات. حقق (AraBERT) تحسينات كبيرة في تحليل المشاعر والتعرف على الكيانات المسماة ومهام معالجة اللغة الطبيعية العربية الأخرى، مما يجعله نموذجاً أساسياً للغة العربية (Antoun et al., 2020).
2. (AraGPT2 2021): قثدم عام (2021)، بتكييف بنية (GPT-2) الخاصة بشركة (OpenAI) للغة العربية. دُرب هذا النموذج التوليدي على مجموعات بيانات خاصة باللغة العربية، متفوقاً في مهام إنشاء النصوص مثل تأليف الشعر ورواية القصص وتوليد الحوار. كما قدم أدوات لغوية إبداعية مصممة خصيصاً للغة العربية (Antoun et al., 2021).
3. (ARBERT and MARBERT 2021): في نفس العام، أُصدر نموذجي (ARBERT) و(MARBERT)، اللذان صمما خصيصاً لمهام معالجة اللغة الطبيعية باللغة العربية. ركزت بيانات التدريب على اللغة العربية الفصحى الحديثة، واللهجات. تناولت هذه النماذج تنوع اللغة العربية وتفوقت على

النماذج السابقة في مهام مثل تحليل المشاعر وتحديد اللهجات (Abdul-Mageed et al., 2021).

4. (AraT5 2022): قدم الباحثون هذا النموذج وهو نسخة من نموذج (T5) من (Google) معدلة للغة العربية. استخدم إطار عمل تحويل النص إلى نص، مما مكنه من أداء مجموعة واسعة من المهام مثل التلخيص والترجمة وتصنيف النصوص. وسعت تنوعاته وقدراته التوليدية نطاق تطبيقات معالجة اللغة الطبيعية باللغة العربية (Nagoudi, 2022).
5. (JASMINE 2023): طُور في عام (2023)، وركز على تعزيز فهم وتوليد اللهجات العربية جنباً إلى جنب مع اللغة العربية الفصحى الحديثة. من خلال دمج البيانات اللهجية في التدريب، حسن الاستجابة لعمليات المحادثة، وخاصة لدعم العملاء والروبوتات في العالم الناطق باللغة العربية (Nagoudi, 2023).
6. (GEMMAr 2024) (General Morphologically-aware Model for Arabic): نموذج قائم على المحولات مصمم خصيصاً للتعامل مع الثراء الصرفي للغة العربية. وعلى عكس النماذج العربية الأخرى، أدرج (GEMMAr) ميزات صرفية صريحة في أثناء عملية التدريب، مثل تصريف الأفعال، وانحرافات الأسماء، والعلامات الإعرابية. وقد أدى هذا النهج إلى تحسين المهام بدرجة أكبر مثل وضع علامات على أجزاء الكلام، والتحليل النحوي، وتحليل المشاعر، وخاصة للنصوص العربية شديدة الانحراف (Chouikhi et al., 2024).
7. (ArabianGPT 2024): يُعد نموذجاً لغوياً كبيراً يركز على اللغة العربية ومصمماً لمعالجة الفجوات في معالجة اللغة الطبيعية العربية، والتي تسبب فيها

هيمنة النماذج التي تركز على اللغة الإنجليزية. وهو جزء من مبادرة (ArabianLLM) الأوسع نطاقًا ويقدم نسختين: (ArabianGPT-0.1B) على نطاق صغير و(ArabianGPT-0.3B) على نطاق متوسط، وكلاهما مصممان خصيصاً للتعقيد اللغوي للغة العربية (Koubaa et al., 2024).

الفصل الرابع: الإطار التطبيقي للدراسة:

في هذا القسم، تُقدم النماذج موضع الدراسة وتوصف كيفية استخدام مجموعة البيانات الناتجة، وخطوات التقييم للحصول على نتائج دقيقة تصف واقع هذه النماذج والوصول إلى مقترحات تُساعد في وضع أسس صحيحة لانتشار تطبيقات الذكاء الاصطناعي في اللغة العربية مع إزالة الآثار السلبية لهذا الهدف.

4.1. اختيار النماذج (ومسوغات اختيارها):

وقع الاختيار على أحدث ثلاثة نماذج صُممت من أكبر شركات التكنولوجيا العالمية، لاختبارها عن طريق تطبيق الدردشة (Chatbot) الخاص بكل منها. اختيرت هذه النسخ لأنها لا تتطلب اشتراكات مادية وبالتالي فهي تستهدف الفئة العادية من المستخدمين وليس مطوري البرامج والمعلوماتيين حيث يوجد نسخ مشابهة من هذه النماذج تستهدف تلك الفئات من حيث الميزات التي تقدمها لدمجها بالأعمال والتطبيقات الأخرى (API Subscribtion, AI Studios, ...) أو تلك المخصصة لمهام أو مجال محدد.

حققت هذه النماذج وفقاً لمطوريها نتائج متطورة على معايير معالجة اللغة الطبيعية المختلفة، مما يجعلها مناسبة لهذه الدراسة المقارنة، فهي متشابهة في البنية ومدربة على مجموعة واسعة من البيانات، مما يسمح باستكشاف نقاط القوة والضعف في كل منها واستكشاف تأثيرات خيارات التصميم المختلفة في الأداء.

4.1.1. (GPT-4o mini OpenAI):

أطلقت (OpenAI) نموذج (GPT-4o mini)، وهو نموذج صغير الأكثر كفاءة من حيث التكلفة. وتتوقع أن يعمل (GPT-4o mini) على توسيع نطاق التطبيقات المبنية بالذكاء الاصطناعي بدرجة أكبر من خلال جعل الذكاء الاصطناعي أقل تكلفة.

يُتيح (GPT-4o mini) مجموعة واسعة من المهام بتكلفة منخفضة وزمن انتقال منخفض، مثل التطبيقات التي تسلسل أو تنسق مكالمات نموذجية متعددة، أو تمرر حجماً كبيراً من السياق إلى النموذج، أو تتفاعل مع العملاء من خلال استجابات نصية سريعة وفي الوقت الفعلي.

اليوم، يدعم (GPT-4o mini) النص والرؤية في واجهة برمجة التطبيقات، مع دعم إدخالات ومخرجات النصوص والصور والفيديو والصوت في المستقبل. وبفضل أداة التجزئة المحسنة المشتركة مع (GPT-4o)، أصبح التعامل مع النصوص غير الإنجليزية أكثر فعالية من حيث التكلفة الآن.

يتفوق (GPT-4o mini) على (GPT-3.5 Turbo) وغيره من النماذج الصغيرة في المعايير الأكاديمية عبر كل من الذكاء النصي والاستدلال متعدد الوسائط، ويدعم نفس مجموعة اللغات مثل (GPT-4o)، كما يظهر أداءً قوياً في استدعاء الوظائف، مما يمكن المطورين من بناء تطبيقات تقوم بجلب البيانات أو اتخاذ إجراءات مع أنظمة خارجية، وتحسين الأداء في السياق الطويل مقارنةً ب (GPT-3.5 Turbo) (OpenAI, 2024).

إحصائيات النموذج (OpenAI, 2024):

Type & Algorithm: Transformer, Generative pre-trained transformer

Category الفئة	Criteria المعيار	Percentage النسبة
Text Evaluation تقييم النص	MMLU	82
	GPQA	53.6
	MATH الرياضيات	70.2
	HumanEval التقييم البشري	87.2
Supported Languages اللغات المدعومة		More than 50
Context length طول السياق		128K
Output token علامات الخرج		16.4 K
DateSet Date تاريخ بيانات التدريب		October 2023 and can access the internet if up-to-date information is needed تشرين الأول 2023 ويمكنه الوصول للانترنت عند الحاجة إلى بيانات أحدث

جدول (1) إحصائيات نموذج GPT-4o mini [1]

[1] تم الحصول على المعلومات الواردة في الجدول من موقع الشركة التي أصدرت النموذج كما ورد في التوثيق أعلاه.

4.1.2. (Gemini 1.5 F Google):

أحدث ابتكارات (Google DeepMind) هو نموذج (Gemini 1.5 Flash)، الذي قُدم في (2024)، على استعداد لإحداث ثورة في عالم التفاعل مع الآلات. يتميز هذا النموذج بأنه متعدد الوسائط يتميز بسرعة وكفاءة وتنوع، مما يمكنه من معالجة وإنشاء النصوص والصور والصوت والفيديو بدقة وطلاقة لا مثيل لها.

إنه نموذج خفيف، مما يعني أنه يتطلب طاقة حوسبة أقل للتشغيل مقارنة بنماذج الذكاء الاصطناعي القوية الأخرى. مما يجعله مثالياً للمهام ذات الحجم الكبير. على الرغم من كونه خفيف، إلا أن (Gemini 1.5 Flash) لا يزال أداة قوية. تتمثل إحدى نقاط القوة الرئيسية في (Gemini 1.5 Flash) في فهمه للسياق الطويل. يمكنه معالجة كميات هائلة من المعلومات، والاحتفاظ بالسياق لتقديم استجابات دقيقة. عموماً، يعد (Gemini 1.5 Flash) إضافة قيمة إلى تطبيقات الذكاء الاصطناعي فهو يسد الفجوة للمستخدمين الذين يحتاجون إلى نموذج ذكاء اصطناعي سريع وبأسعار معقولة ومتعددة الاستخدامات (Google Deemind, 2024).

إحصائيات النموذج (ContextAI):

Type & Algorithm: Transformer, Generative pre-trained transformer

Category الفئة	Criteria المعيار	Percentage النسبة
Text Evaluation تقييم النص	MMLU	78.9
	GPQA	39
	MATH الرياضيات	54.9
	HumanEval التقييم البشري	74.3
Supported Languages اللغات المدعومة		More than 40
Context length طول السياق		1M
Output token علامات الخرج		8192
DateSet Date تاريخ بيانات التدريب		November 2023 and can access the internet if up-to-date information is needed تشرين الثاني 2023 ويمكنه الوصول للانترنت عند الحاجة إلى بيانات أحدث

جدول (2) إحصائيات نموذج Gemini 1.5 F [2]

[2] تم الحصول على المعلومات الواردة في الجدول من الموقع الذي ورد في التوثيق أعلاه.

4.1.3. (LLAMA 3.1-70B Instruct Meta AI):

نموذج من نماذج اللغات الكبيرة التوليدية المدربة مسبقاً متعددة اللغات والمضبوطة وفقاً لتعليمات بحجم يبلغ 70B. طورته (Meta) ودربته لحالات استخدام الحوار متعدد اللغات وتفوق على العديد من نماذج الدردشة المفتوحة المصدر والمغلقة المتاحة وفقاً للمعايير الشائعة.

وهو أيضاً نموذج لغوي انحداري تلقائي يستخدم بنية محول محسّنة، يستخدم الضبط الدقيق الخاضع للإشراف والتعلم التعزيزي مع ردود الفعل البشرية للتوافق مع التفضيلات البشرية للمساعدة والسلامة (Hugging Face, 2024).

تشير لاحقة "Instruct" إلى أن النموذج ضُبط بدقة لمهام اتباع التعليمات. وهذا يعني أن النموذج دُرب على مجموعة بيانات تؤكد على اتباع التعليمات وتوليد نص ذي صلة ومتنوع ومتماسك ودقيق في الاستجابة لمطالبات أو مهام محددة؛ ونتيجة لذلك، فإن النموذج قادر على فهم التعليمات والاستجابة لها، مما يجعله مناسباً لمجموعة واسعة من التطبيقات.

عموماً، يعد أداة قوية لديها القدرة على إحداث ثورة في الطريقة التي نتفاعل بها مع نماذج اللغة. تجعلها قدراتها المتقدمة على فهم اللغة وقدراتها على متابعة التعليمات وحجمها الهائل حلاً جذاباً لاستخدامها في العديد من التطبيقات، من المساعدين الافتراضيين إلى توليد المحتوى.

إحصائيات النموذج (Hugging Face, 2024):

Type & Algorithm: Transformer, Generative pre-trained transformer

Category الفئة	Criteria المعيار	Percentage النسبة
Text Evaluation تقييم النص	MMLU	83.6
	GPQA	46.7
	MATH الرياضيات	68
	HumanEval التقييم البشري	80.5
Supported Languages اللغات المدعومة		8
Context length طول السياق		128K tokens
Output token علامات الخرج		2048
DateSet Date تاريخ بيانات التدريب		December 2023 and can access the internet if up-to-date information is needed كانون الأول 2023 ويمكنه الوصول للانترنت عند الحاجة إلى بيانات أحدث

جدول (3) إحصائيات نموذج LLAMA 3.1-70B Instruct [3]

[3] تم الحصول على المعلومات الواردة في الجدول من الموقع الذي ورد في التوثيق أعلاه.

4.2. جمع البيانات:

4.2.1. تصميم المطالبة (Prompt Design):

إن المطالبة هي نهج يتبعه المستخدمون للتفاعل مع نماذج التعلم اللغوي (White et al., 2023). وهو يعمل كوسيلة أساسية للتواصل مع هذه النماذج، وبفعالية كلغة إدخال صممت نماذج التعلم اللغوي لتفسيرها والاستجابة لها، وترتبط فعالية الناتج بدرجة أكبر بجودة وبنية التوجيه المدخل. وتؤكد هذه العلاقة على الدور الحاسم الذي تلعبه هندسة التوجيه في تحسين أداء نماذج التعلم اللغوي وأهمية الإخراج.

وجدت الباحثة أن المدخلات الخاصة بنماذج اللغة الكبيرة أو "المطالبات"، تشكل أهمية بالغة لقدرتها على إكمال مجموعة متنوعة من المهام. أتاحت نماذج اللغة الكبيرة المدربة مسبقاً الوصول إليها من قبل الأشخاص الذين لا يحتاجون إلى تدريب على البرمجة. حيث شهدت اللغة العربية في الأعوام الأخيرة انتشاراً في المحتوى الذي يعكس نتائج تجارب الأشخاص العاديين مع النماذج لتوليد المحتوى والمزيد.

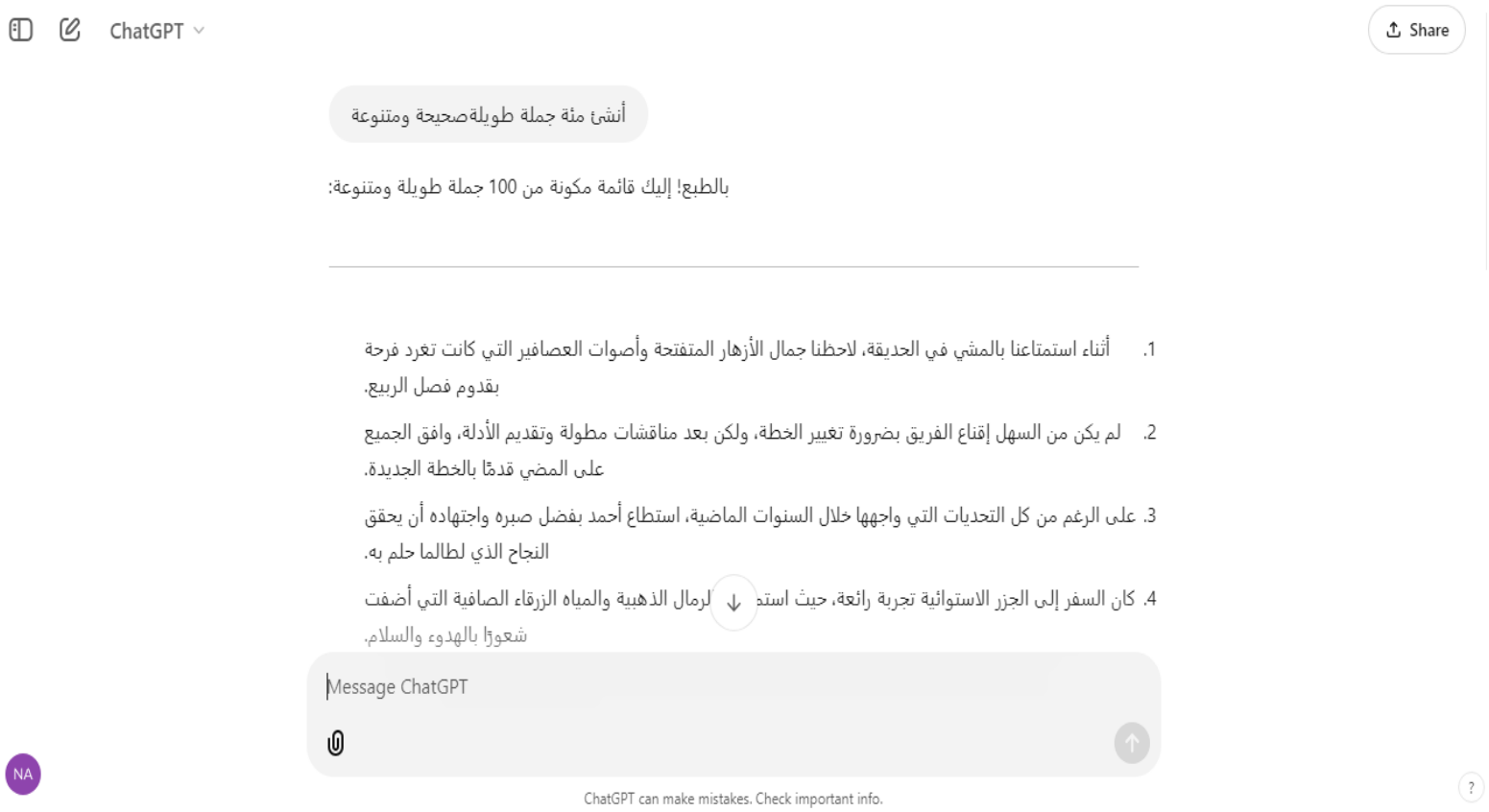
وتأتي المطالبات بأشكال مختلفة، وحتى بالنسبة لنفس المطالبة فإن طريقة التعبير المختلفة لدى الأفراد المختلفين تؤدي إلى استجابات مختلفة، حيث تعمل المطالبات على تحيز النموذج مباشرة نحو توليد المخرجات المرغوبة. يُذكر أربعة أنماط من المطالبات التي يستخدمها النماذج عند إكمال المهام: الإدخال البسيط (المشكلة، الحل)، الدعم العاطفي (رجاءاً قدم حل للمشكلة الآتية)، لعب الدور (أنت باحث قم بإيجاد حل للأسئلة الآتية)، مطلب النتيجة (رجاءاً أجب عن السؤال الآتي وضمن الإجابة بعد السؤال) (Zhuo, 2024).

جُربت الطرائق الثلاث الأولى التي وردت في المرجع السابق على النماذج الثلاثة المختارة في هذا البحث، ولم تطبق الطريقة الأخيرة "مطلب النتيجة"؛ لأنها تعتمد مبدأ السؤال والجواب وهي لا تتوافق مع العينة المطلوبة في هذا البحث، فالباحثة لا تريد توليد إجابات على أسئلة، بل تريد التركيز على توليد المحتوى ولذلك اقتصرت على الطرائق الثلاث الأولى:

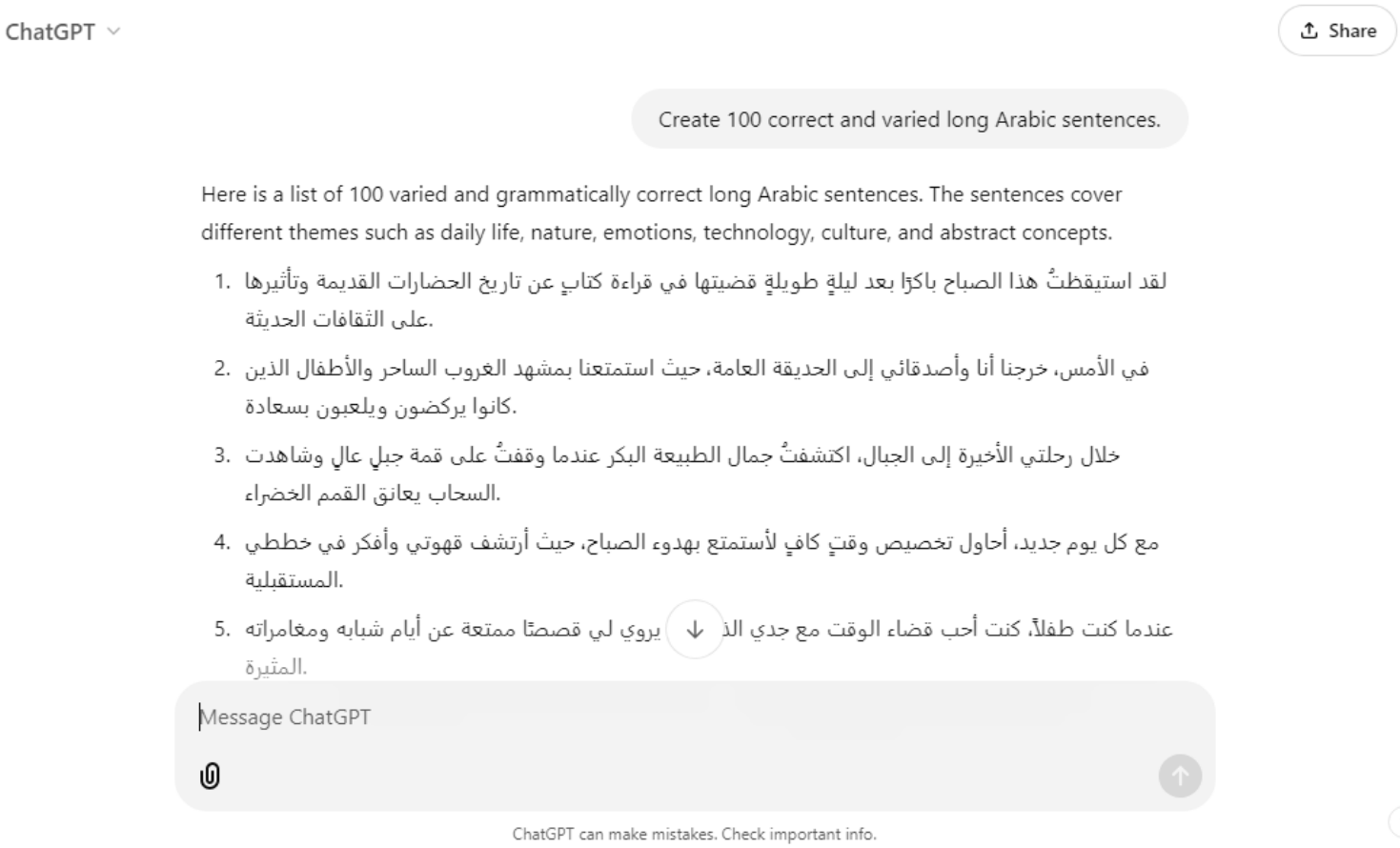
1. (GPT-4o mini OpenAI) (عن طريق ChatGPT Chatbot):

- طريقة الإدخال البسيط:

➢ المطالبة باللغة العربية:



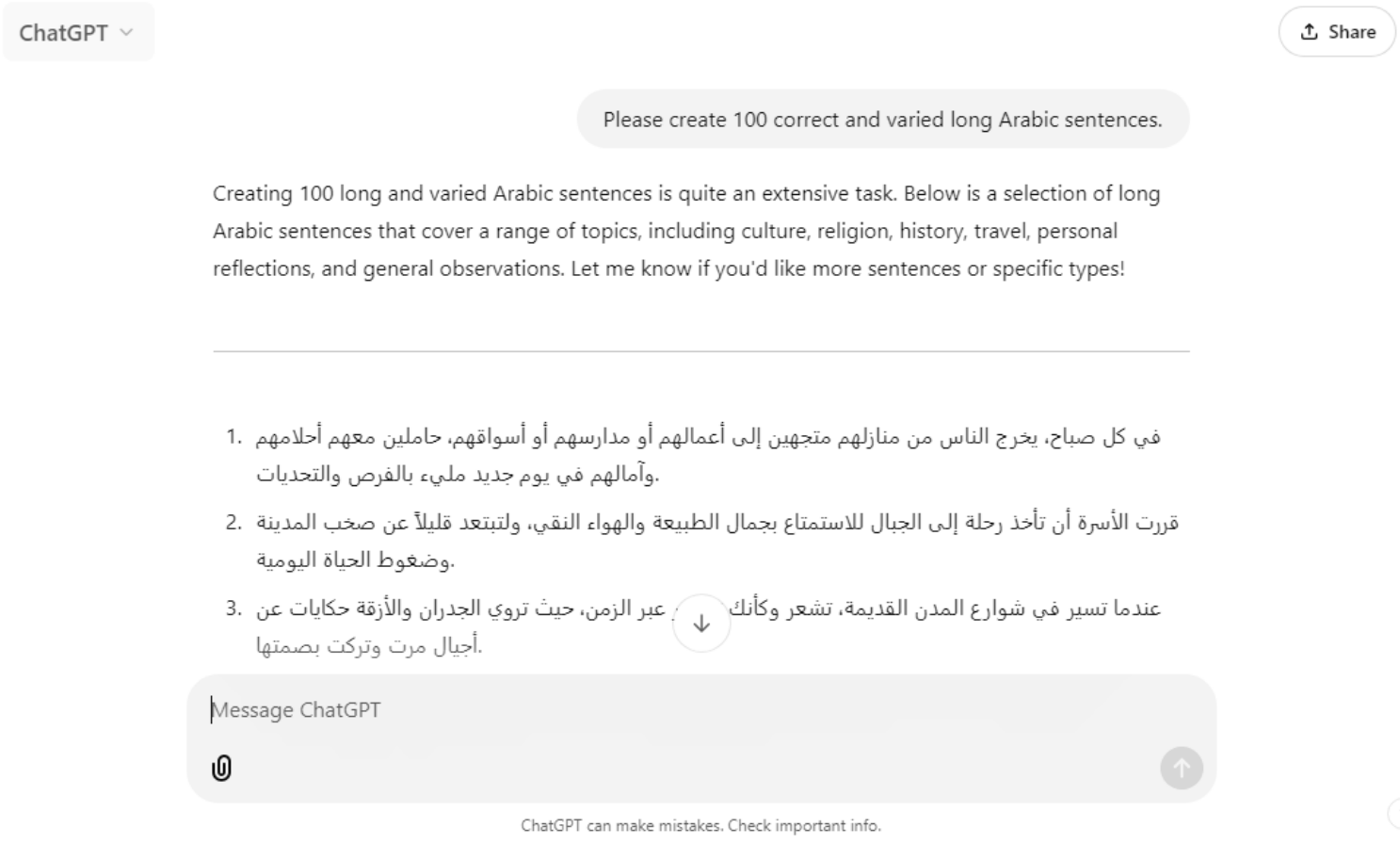
➢ المطالبة باللغة الإنكليزية:



- طريقة الدعم العاطفي:

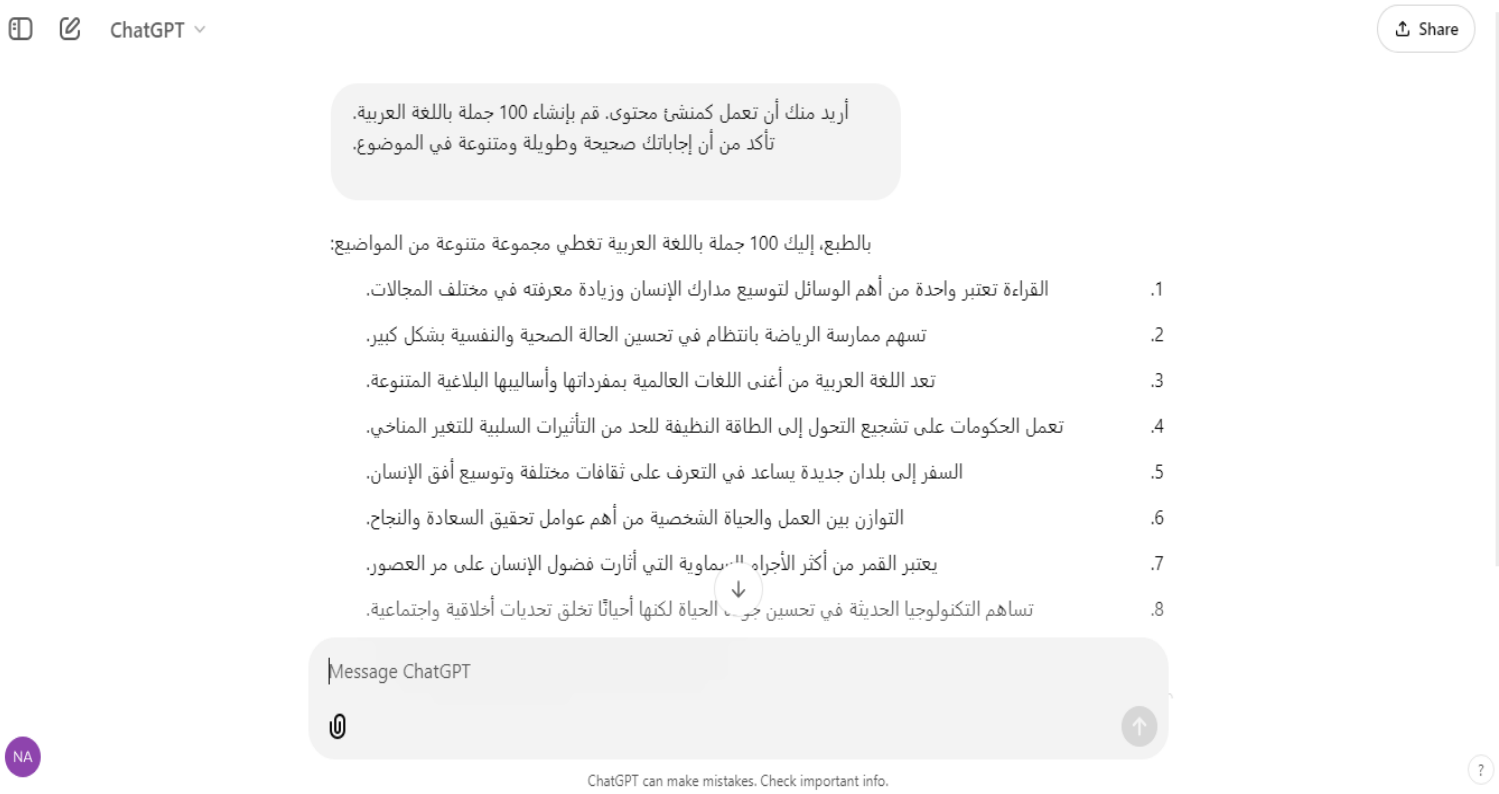
➢ المطالبة باللغة العربية:

➢ المطالبة باللغة الإنكليزية:

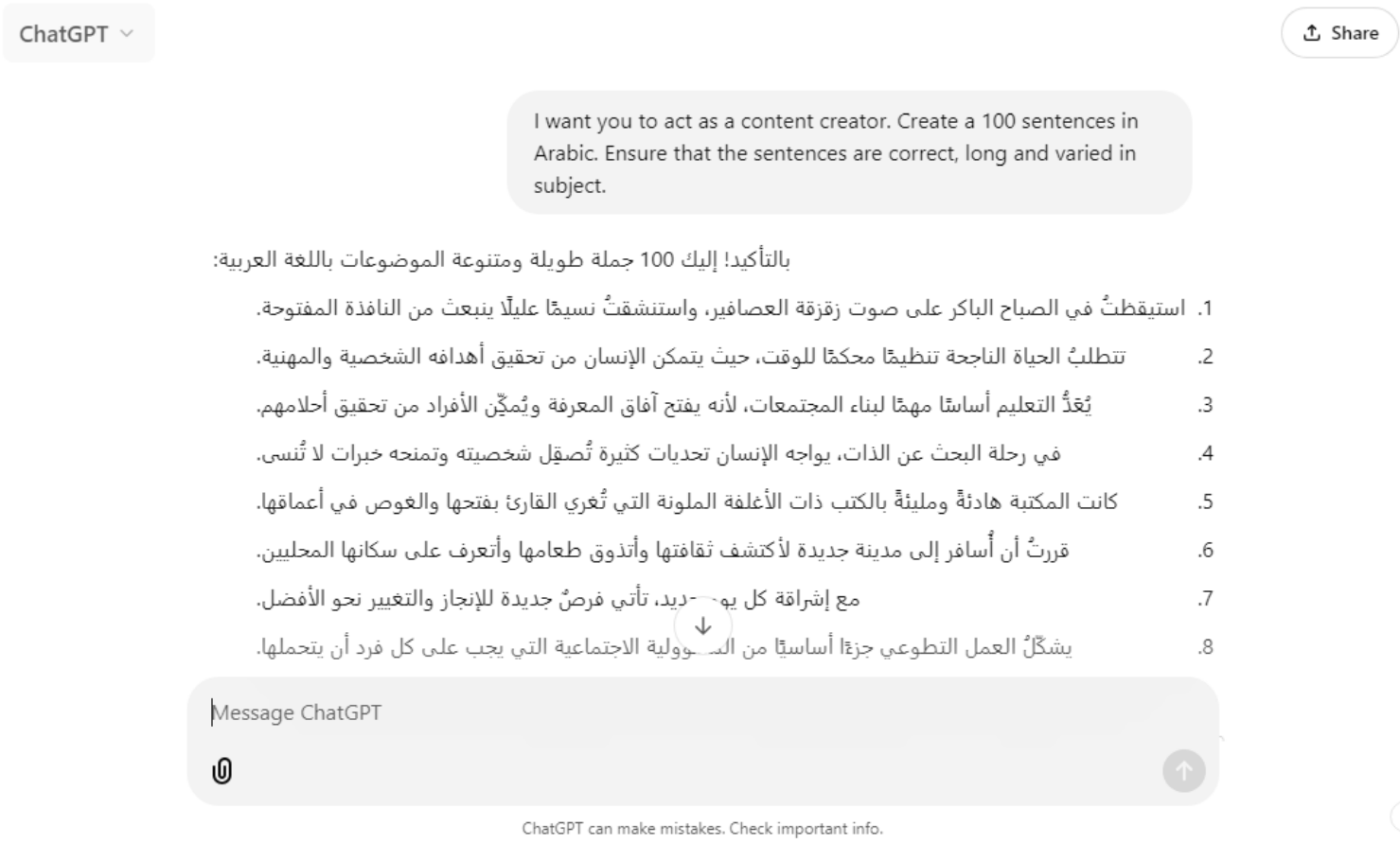


- طريقة لعب الدور:

➢ المطالبة باللغة العربية:



➢ المطالبة باللغة الإنكليزية:



2. (Gemini 1.5 F Google) (عن طريق Gemini Chatbot):

- طريقة الإدخال البسيط:

➢ المطالبة باللغة العربية:

سُعي لتخصيص المطالبة بدرجة أكبر للحصول على استجابة أفضل من النموذج تطابق هدف العينة، ومحاولة التأكيد على عملية توليد المحتوى كانت الاستجابة ذاتها تقريباً مع تقديم بعض الأمثلة من النموذج

- **التحدث العام:** لإقناع الآخرين أو إثارة تفكيرهم.

إليك بعض الأمثلة:

جمل وصفية:

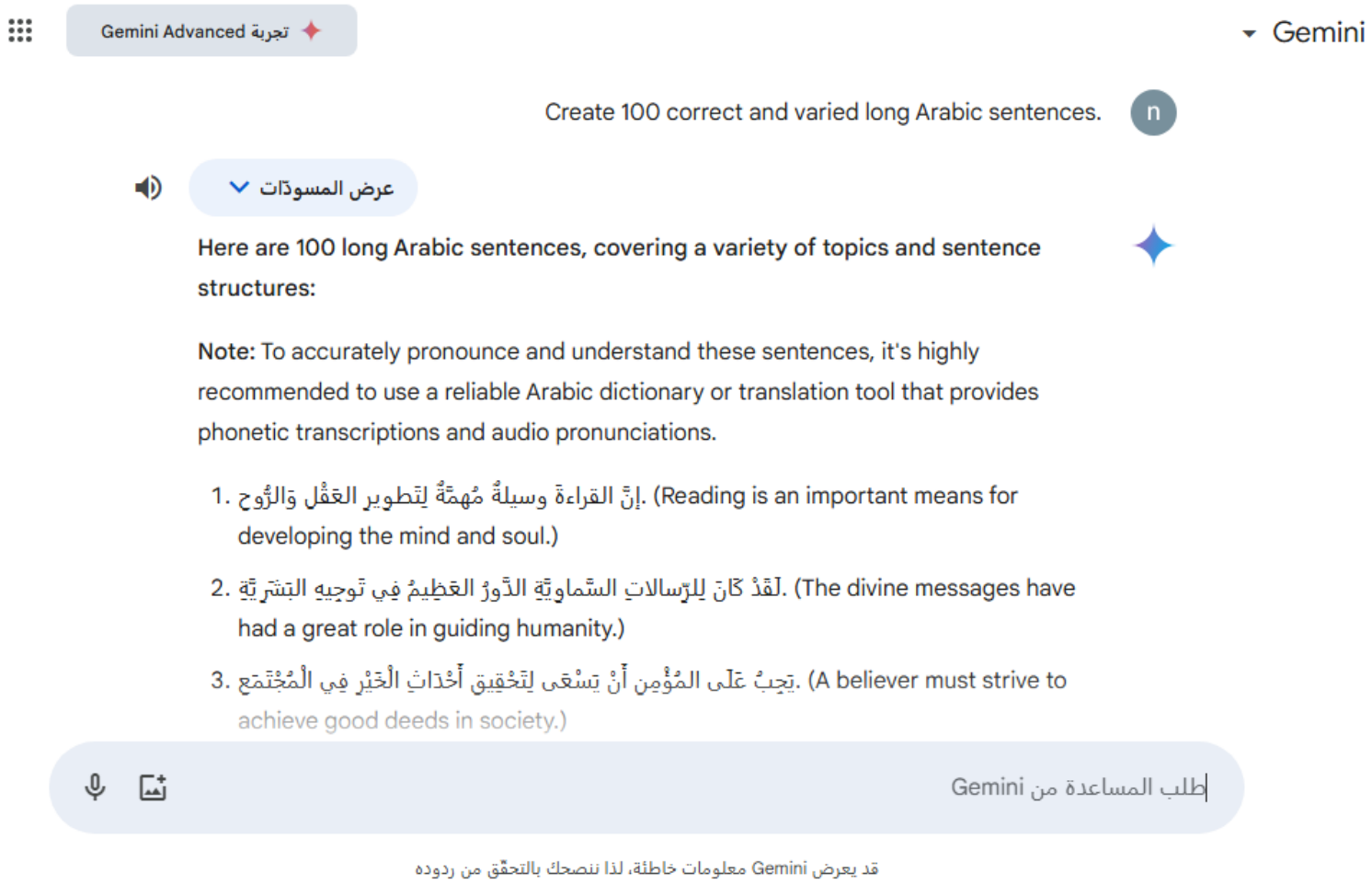
- كانت الغابة المظلمة كقلب الإنسان، مليئة بالألغاز والأسرار، حيث تتشابك الأشجار الضخمة وكأنها تحاول حجب ضوء الشمس، وتصدر أصوات غريبة تشبه أنين الأرواح، مما يثير في النفس مشاعر الخوف والرهبة.
- بحر الشفق يمتد كلوحة فنية رائعة، حيث تتداخل ألوان البرتقالي والأحمر والبنفسجي، وتنعكس أشعة الشمس الذهبية على الأمواج المتلاطمة، مما يخلق منظرًا ساحرًا يبعث على الهدوء والسكينة.

جمل حوارية:

- قال الشاب بحماس: "إنني أؤمن بأن العلم هو مفتاح المستقبل، وأننا بواسطته يمكننا حل جميع مشاكل العالم، ولكن علينا أن نكون حذرين من استخدامه في أغراض ضارة، فالقوة العظمى تأتي مع مسؤولية عظمى."
- بدت العجوز بحكمة: "الحياة كالبحر، مليئة بالموجات الصاعدة والهابطة، فلا تيأس عندما تواجه

طلب المساعدة من Gemini

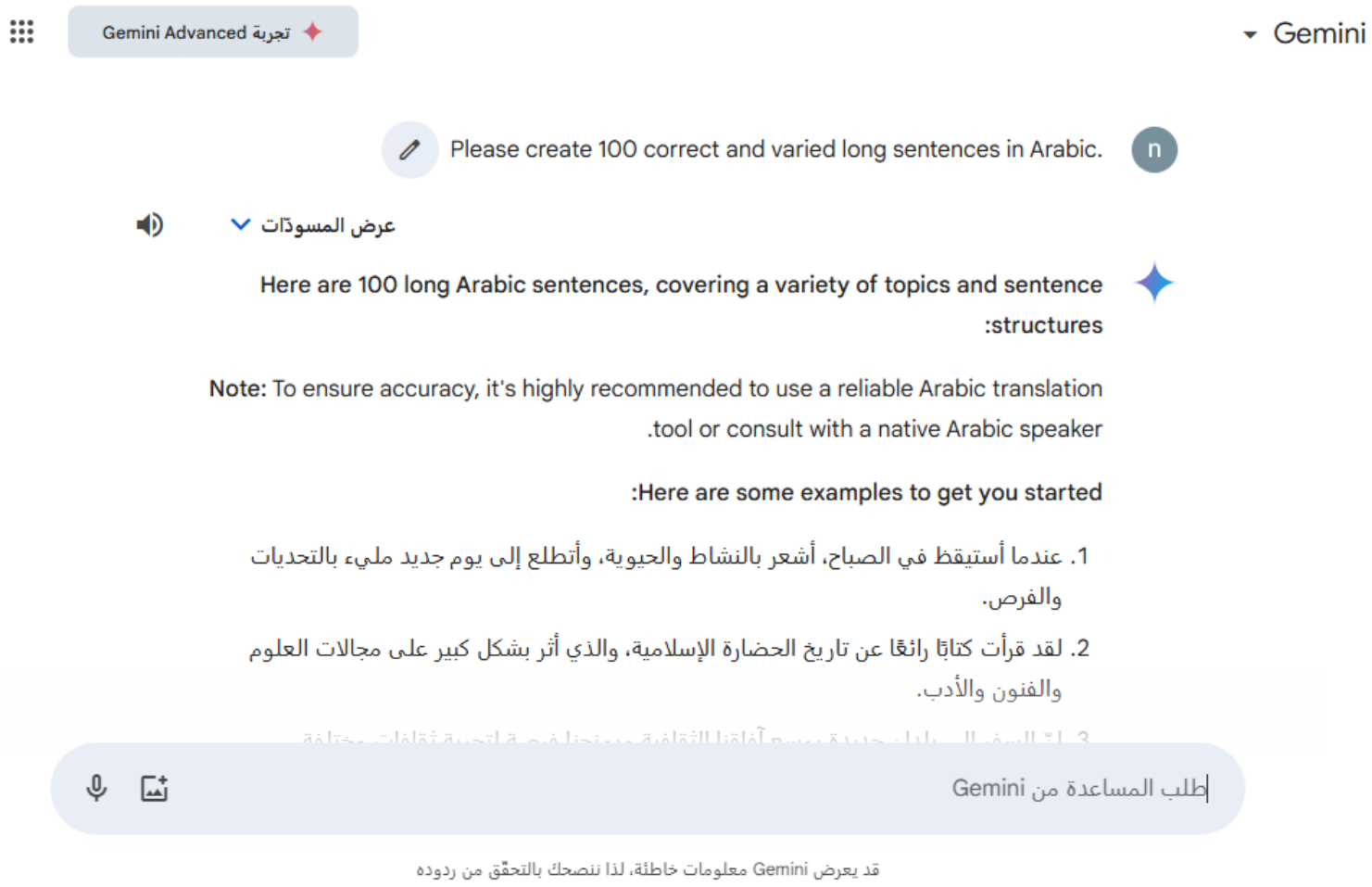
➢ المطالبة باللغة الإنكليزية:



- طريقة الدعم العاطفي:

➢ المطالبة باللغة العربية:

➢ المطالبة باللغة الإنكليزية:



- طريقة لعب الدور :

➢ المطالبة باللغة العربية:

جمل تعبيرية

- "تحتفي روح المدينة بألوانها الزاهية، وكأنها لوحة فنية تتجدد كل يوم."
- "في أعماق أعينها بحر من الأسرار، وفي ابتسامتها ربيع يزهو."
- "صوت الرياح يحكي حكايات العصور القديمة، وأشجار النخيل تشهد على عراقة الأرض."
- "الكتاب هو صديقي الدائم، يرافقني في كل رحلة وأستكشف معه عوالم جديدة."

جمل فلسفية

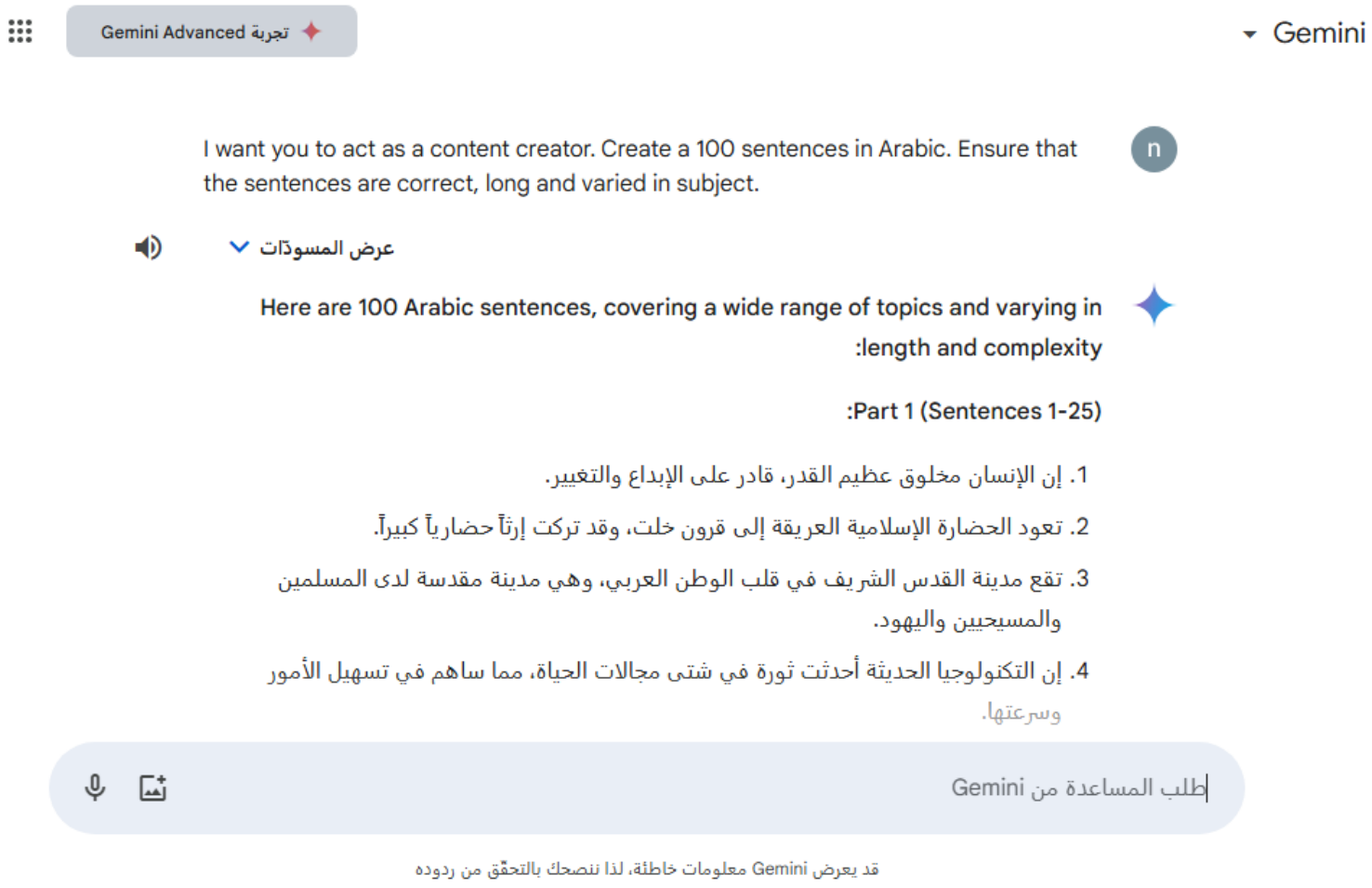
- "الحياة هي لغز نسعى جاهدين لفك شفرته، وكل إنسان يجد إجابته الخاصة."
- "السعادة ليست هدفًا نصل إليه، بل هي رحلة نستمتع بكل لحظة فيها."
- "الحقيقة مطلقة، ولكن تفسيراتنا لها متعددة."
- "الوقت كالسيف، إن لم تقطعه قطعك."

جمل أدبية

➢ المطالبة باللغة الإنكليزية:

رفض النموذج إنشاء أي من الجمل وحتى عند محاولة التأكيد لم ينتج أي محتوى.

وعند المحاولة مرة جديدة في محادثة جديدة قام بإنشاء عينة جيدة تقارب الهدف.



3. (LLAMA 3.1-70B Instruct Meta) (عن طريق Hugging Face Chatbot):

- طريقة الإدخال البسيط:

➢ المطالبة باللغة العربية:

المحاولة الأولى:

أنشئ مئة جملة طويلة صحيحة ومتنوعة

1. ذهب محمد إلى المدرسة صباحاً.
2. ذهبت أمي للمدرسة صباحاً.
3. ذهب جدي إلى الحديقة صباحاً.
4. ذهبت أختي إلى المكتبة صباحاً.
5. ذهبت أمي إلى السوق.
6. ذهب جدي إلى المسجد.
7. ذهب محمد إلى الحديقة.
8. ذهبت أختك إلى المدرسة.
9. ذهب أخي إلى المكتبة.
10. ذهبت أختي إلى السوق.

Tools (0)

Ask anything

Model: meta-llama/Meta-Llama-3.1-70B-Instruct · Generated content may be inaccurate or false.

Share this conversation

المحاولة الثانية:

➢ المطالبة باللغة الإنكليزية:

- طريقة الدعم العاطفي:

➢ المطالبة باللغة العربية:

➢ المطالبة باللغة الإنكليزية:

- طريقة لعب الدور:

➢ المطالبة باللغة العربية:

➢ المطالبة باللغة الإنكليزية:

I want you to act as a content creator. Create a 100 sentences in Arabic. Ensure that the sentences are correct, long and varied in subject.

Here are 100 sentences in Arabic, covering a range of subjects:

1. كانت مدينة بغداد في العصور الوسطى مركزاً رئيسياً للثقافة والتعليم والفلسفة.
2. يعتبر القاهرة مدينة عالمية ومحورًا اقتصاديًا وثقافيًا في الشرق الأوسط.
3. تتميز الأهرامات المصرية بجمالها وغرابتها المعمارية.
4. كانت مصر القديمة حضارة متقدمة للغاية في مجالات الهندسة المعمارية والفلك والطب.
5. يعتبر نهر النيل مصدرًا رئيسيًا للمياه في مصر والبلدان المحيطة بها.
6. تضم مصر مجموعة واسعة من الآثار التي تعود إلى العصور القديمة.
7. يعتبر الفن الإسلامي من الفنون الجميلة التي تتميز بالزخارف والنقوش.
8. تأتي الفواكهة بجميع أنواعها وألوانها إلى الأسواق المصرية.
9. يمكن للزوار استكشاف تاريخ مصر من خلال المتاحف والمuseums الموجودة في جميع أنحاء البلاد.
10. يعتبر السوق التقليدي في القاهرة وجهة مثالية للتسوق والاستكشاف.

Tools (0) Stop generating

Ask anything

Model: meta-llama/Meta-Llama-3.1-70B-Instruct · Generated content may be inaccurate or false. Share this conversation

نتائج محاولات المطالبة:

1. يُستنتج من التجارب السابقة لمحاولة تصميم المطالبة بأن استجابات نموذج (GPT-4o mini) كانت متماثلة للمطالبات باللغة الإنكليزية والعربية في طرائق تصميم المطالبة (الإدخال البسيط والدعم العاطفي ولعب الدور) ومتنوعة محققة لهدف إنشاء عينة مناسبة للدراسة؛ مما يدل على تدريب النموذج على طيف بيانات واسع باللغة العربية وتعرضه لعمليات ضبط مخصصة للغة العربية.

2. يُستنتج من التجارب السابقة لمحاولة تصميم المطالبة بأن استجابات نموذج (Gemini 1.5 F) في حالة المطالبة باللغة الإنكليزية أكثر دقة وقرباً للهدف من المطالبة باللغة العربية في طريقتي تصميم المطالبة بالإدخال البسيط والدعم العاطفي. أما طريقة لعب الدور، فكانت النتيجة أفضل من طريقتي الإدخال البسيط والدعم العاطفي في حالة المطالبة باللغة العربية، بينما في حالة المطالبة باللغة الإنكليزية احتاج النموذج لمحاولات عدة لإنتاج عينة مرغوبة.

3. يُستنتج من التجارب السابقة لمحاولة تصميم المطالبة بأن استجابات نموذج (LLAMA 3.1-70B Instruct) في حالة المطالبة باللغة العربية في طريقة تصميم المطالبة بالإدخال البسيط، أنتجت المحاولة الأولى جملاً قصيرة وغير متنوعة الموضوع، والمحاولة الثانية أنتجت جملاً متطابقة وغير مناسبة، أما حالة المطالبة باللغة الإنكليزية كانت نتيجة إنشاء العينة بأن الجمل متشابهة من حيث المفردات والبنية. وكانت استجابات نموذج (LLAMA 3.1-70B Instruct) في حالة المطالبة باللغة العربية واللغة الإنكليزية في طريقتي الدعم العاطفي ولعب الدور متشابهة نوعاً ما، فإن الجمل المولدة تكررت فيها المفردات والأنماط والمواضيع، ولكن في طريقة لعب الدور في حالة المطالبة باللغة الإنكليزية احتوت بعض الجمل على كلمات (أو

رموز) من لغات أخرى. وأفضل نتائجه (على الرغم من أن الفرق بسيط) كانت بطريقة الدعم العاطفي بالمطالبة باللغة العربية ومن ثم طريقة لعب الدور بالمطالبة باللغة الإنكليزية.

4. يُستنتج أن استجابات نموذج (LLAMA 3.1-70B Instruct) لم تكن كما يجب، وعند مراجعة بطاقة النموذج يُلحظ أنه لا يدعم اللغة العربية، وذلك لا يعني بأنه لا يفهم اللغة العربية أو لا يستطع توليد المحتوى باللغة العربية، بل بأنه لم يدرب على كمية كبيرة من بيانات النص بها، أو لم يُضبط لأداء جيد للمهام بالعربية.

5. ومما سبق يُستنتج أن أفضل تصميم للمطالبة هو اعتماد مزيج بين طريقتي الدعم العاطفي ولعب الدور واعتماد المطالبة باللغة الإنكليزية لإنتاج محتوى باللغة العربية مع إضافة تفاصيل مخصصة قدر الإمكان للحصول على أفضل عينة للدراسة من النماذج الثلاثة مع مراعاة توليد العينة على مراحل (ضمن محادثات جديدة) لتقليل فرص التكرار الموضوعي واللغوي.

وجُرب بمسودات عدة من المطالبات السابقة حتى الوصول إلى شكلها الآتي المتوقع أن ينتج أفضل عينة متنوعة ويقلل فيها التكرار قدر الإمكان وفق التحليل البشري.

المطالبة النهائية:

Act as a content creator and generate x long and useful sentences in Arabic. Please ensure that your responses use different topics, diverse range of sentence structures and vocabulary. Make sure that the generated sentences are long and at different levels of difficulty. Do not add titles or categorize them, do not add any translation or pronunciations to them and do not rely on repetitive patterns between sentences.

4.2.2. توليد العينة:

- أُعطيت المطالبة السابقة ل Chatbot الخاصة بكل نموذج خاضع للدراسة (LLAMA 3.1-70B Instruct – GPT-4o mini – Gemini 1.5 F).
- أُعطيت (4) مرات بحيث طلب في كل مرة توليد (25) جملة (بدل x).
- أُضيفت هذه الجمل لملف خارجي لكي تخضع للتقييم في الخطوات اللاحقة.


Sample.xlsx

4.2.3. تحليل العينة:

4.2.3.1. تحليل العينة من حيث تنوع المحتوى:

يُقاس عدد المواضيع التي ظهرت في مجموعات العينة (الناتجة عن كل محادثة) وفيها مجموعةً. يقارن الجدول الآتي أداء النماذج الثلاثة عبر أربع مجموعات:

Model Comparison from Subjects Varity					
Model	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Total**
GPT-4o mini	20	17	15	19	34
	80%	68%	60%	76%	
Gemini 1.5 F	17	14	16	19	30
	68%	56%	64%	76%	
LLAMA 3.1-70B Instruct	14	18	10	15	29
	56%	72%	40%	60%	

جدول (4) النتائج من حيث تنوع المحتوى

4.2.3.2. تحليل العينة من حيث تحليل المشاعر:

يُقاس عدد المواضيع التي ظهرت في مجموعات العينة (الناتجة عن كل محادثة) وفيها مجموعةً وتُصنف حسب العاطفة التي تطغى عليها وفق ما يراه النموذج نفسه لفئات تحليل المشاعر الثلاث: إيجابية وسلبية ومحايدة. يقارن الجدول الآتي أداء النماذج الثلاثة عبر أربع مجموعات:

Sentiment Analysis	Positive					Negative					Neutral				
Model	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total
GPT-4o mini	20	13	22	21	76	0	2	1	0	3	5	10	2	4	21
	80%	52%	88%	84%		0%	8%	4%	0%		20%	40%	8%	16%	
Gemini 1.5 F	20	23	18	23	84	5	0	6	2	13	0	2	1	0	3
	80%	92%	72%	92%		20%	0%	24%	8%		0%	8%	4%	0%	
LLAMA 3.1-70B Instruct	18	21	18	21	78	2	0	3	4	9	5	4	4	0	13
	72%	84%	72%	84%		8%	0%	12%	16%		20%	16%	16%	0%	

جدول (5) النتائج من حيث تحليل المشاعر

4.2.3.3. تحليل العينة من حيث البنى النحوية التركيبية (الجملة الابتدائية):

يُقاس عدد البنى النحوية التي ظهرت في مجموعات العينة (الناتجة عن كل محادثة) وفيها مجموعةً. يقارن الجدول الآتي أداء النماذج الثلاثة عبر أربع مجموعات:

(Starting) Sentence Structure	Verbial					Nominal					Adverbial				
Model	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total
GPT-4o mini	11	17	5	14	47	6	3	7	6	22	4	1	6	1	12
	44%	68%	20%	56%		24%	12%	28%	24%		16%	4%	24%	4%	
Gemini 1.5 F	11	7	4	1	23	10	12	20	22	64	0	0	0	0	0
	44%	28%	16%	4%		40%	48%	80%	88%		0%	0%	0%	0%	
LLAMA 3.1-70B Instruct	22	17	17	24	80	0	7	3	0	10	0	0	0	0	0
	88%	68%	68%	96%		0%	28%	12%	0%		0%	0%	0%	0%	
(Starting) Sentence Structure	Conditional					Prepositional					Verbial preceeded by a partical				
Model	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total	Group 1	Group 2	Group 3	Group 4	Total
GPT-4o mini	0	0	1	0	1	4	4	4	2	14	0	0	2	2	4
	0%	0%	4%	0%		16%	16%	16%	8%		0%	0%	8%	8%	
Gemini 1.5 F	0	0	0	0	0	4	6	1	2	13	0	0	0	0	0
	0%	0%	0%	0%		16%	24%	4%	8%		0%	0%	0%	0%	
LLAMA 3.1-70B Instruct	0	0	0	0	0	3	1	5	1	10	0	0	0	0	0
	0%	0%	0%	0%		12%	4%	20%	4%		0%	0%	0%	0%	

جدول (6) النتائج من حيث البنى النحوية التركيبية

4.3. الإجابة عن أسئلة الدراسة وعرض النتائج ومناقشتها:

تتحدد مشكلة البحث بالإجابة عن الأسئلة الآتية:

- **ما مدى فعالية خوارزميات الذكاء الاصطناعي المدربة في توليد محتوى عربي متنوع؟**

<u>إجابة السؤال الأول:</u>

بعد النظر إلى ما ورد في العينة وتحليلها من حيث تنوع المحتوى، يُستنتج الآتي:

- **النموذج الأول (GPT-4o mini):**

1. يحقق أعلى نسبة تنوع المواضيع في النموذج في المجموعة (1) والمجموعة (4)، فكانت (80%) و(76%) على التوالي.

2. كانت النتائج بنسب عالية في جميع المجموعات، وحتى على مستوى العينة كلها (34%) مما يعكس قدرة كبيرة على إنتاج محتوىً متنوعاً قوياً للمواضيع.

- **النموذج الثاني (Gemini 1.5 F):**

1. يظهر قوة خاصة في المجموعة (3) بنسبة (64%)، متفوقاً على كل من (Gemini 1.5 F) و(LLAMA 3.1-70B Instruct) فيها.

2. أداؤه في المجموعة (4) يطابق (GPT-4o mini) بنسبة (76%)، مما يُظهر إمكانات توليد محتوى قوي تنوع المواضيع على الرغم من تراجع أداؤه مقابله عموماً.

- **النموذج الثالث (LLAMA 3.1-70B Instruct):**

1. يبرز فقط في المجموعة (2) بنسبة تنوع (72%)، مما يشير إلى التخصص المحتمل في مواضيع أو سيناريوهات محددة.

2. يشير أدائه الأضعف في أماكن أخرى إلى الافتقار إلى القدرة على التكيف عبر توليد محتوى في مجالات موضوعية متنوعة.

معيار مرجعي:

الرقم	المعيار
1	كلما ازدادت النسبة في المجموعات الفرعية/المجموعة الكاملة كلما كان أداء النموذج أفضل في توليد محتوى متنوع

جدول (7) معيار مرجعي من حيث تنوع المحتوى

(ملخص نتيجة 1):

يُستنتج أن نموذج (GPT-4o mini) تفوق في توليد محتوى متنوع الموضوع على النموذجين الآخرين. كما يُستنتج أن نموذجي (LLAMA 3.1-70B Instruct) ونموذج (Gemini 1.5 F) متقاربين في الأداء من حيث توليد محتوى متنوع الموضوع.

بعد النظر إلى ما ورد في العينة وتحليلها من حيث تحليل المشاعر، يُستنتج الآتي:

- **النموذج الأول (GPT-4o mini):**

1. حقق نسب عالية في توليد الاستجابات الإيجابية، (80%-88%) عبر المجموعات، (76%) للعينة كاملة.

2. أداء توليد المحتوى ذو العاطفة الإيجابية متسق نوعاً ما عبر جميع المجموعات الفرعية عدا المجموعة الثانية.

3. الأداء العالي في توليد الاستجابات الإيجابية يعني ضعف التنوع في الاستجابات من الأنواع الأخرى.

4. أداء ضعيف خاصة في توليد الاستجابات السلبية، (0%–8%) عبر المجموعات، (3%) للعينة كاملة.

5. تظهر صعوبة لدى النموذج بالتعرف على المشاعر السلبية أو توليدها، حيث هذه العاطفة غائبة تماماً في محتوى المجموعات (1, 4).

6. أداء معتدل في توليد الاستجابات المحايدة، (8%–40%) عبر المجموعات، (21%) للعينة كاملة.

7. أداء توليد المحتوى ذو العاطفة المحايدة غير متسق عبر جميع المجموعات الفرعية.

- **النموذج الثاني (Gemini 1.5 F):**

1. حقق نسب نتائج كبيرة في توليد الاستجابات الإيجابية، (72%–92%) عبر المجموعات، (84%) للعينة كاملة.

2. الأداء العالي في توليد الاستجابات الإيجابية يعني ضعف التنوع في الاستجابات من الأنواع الأخرى.

3. يُظهر بعض التحديات في توليد الاستجابات السلبية، (0%–24%) عبر المجموعات، (13%) للعينة كاملة.

4. العاطفة السلبية غائبة تماماً في المجموعة الثانية.

5. أداء ضعيف خاصة في توليد الاستجابات المحايدة، (0%–8%) عبر المجموعات، (3%) للعينة كاملة.

6. تظهر صعوبة لدى النموذج بالتعرف على المشاعر المحايدة أو توليدها، حيث هذه العاطفة غائبة تماماً في محتوى المجموعات (1, 4).

- **النموذج الثالث (LLAMA 3.1-70B Instruct):**

1. أداء قوي في توليد الاستجابات الإيجابية، (72%-84%) عبر المجموعات، (78%) للعينة كاملة.

2. أداء توليد المحتوى ذو العاطفة الإيجابية متسق نوعاً ما عبر جميع المجموعات الفرعية.

3. الأداء العالي في توليد الاستجابات الإيجابية يعني ضعف التنوع في الاستجابات من الأنواع الأخرى.

4. أداء معتدل يميل للضعف في توليد الاستجابات السلبية، (0%-16%) عبر المجموعات، (9%) للعينة كاملة.

5. تظهر صعوبة لدى النموذج بالتعرف على المشاعر السلبية أو توليدها، حيث هذه العاطفة غائبة تماماً في محتوى المجموعة (2).

6. أداء معتدل يميل للضعف في توليد الاستجابات المحايدة، (0%-20%) عبر المجموعات، (13%) للعينة كاملة.

7. تظهر صعوبة لدى النموذج بالتعرف على المشاعر المحايدة أو توليدها، حيث هذه العاطفة غائبة تماماً في محتوى المجموعة (4).

معيار مرجعي:

الرقم	المعيار
1	كلما ازدادت النسبة في توليد الاستجابات من نوع واحد من المشاعر في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أسوأ في توليد محتوى متنوع المشاعر
2	كلما انخفضت النسبة في توليد الاستجابات من نوع واحد من المشاعر في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أسوأ في توليد محتوى متنوع المشاعر
3	عندما تتقارب النسب في توليد الاستجابات من أنواع المشاعر الثلاثة في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أفضل في توليد محتوى متنوع المشاعر

جدول (8) معيار مرجعي من حيث تحليل المشاعر

(ملخص نتيجة 2):

يُستنتج أن نموذج (LLAMA 3.1-70B Instruct) تفوق في توليد محتوى متنوع المشاعر. على الرغم من طغيان المحتوى ذو المشاعر الإيجابية مع وضوح نسب المحتوى ذو المشاعر السلبية والمحايدة بالمقارنة مع النموذجين الآخرين في نفس الفئات. ولا بد من ذكر أن أداء نموذج (GPT-4o mini) يتفوق على أداء نموذج (LLAMA 3.1-70B Instruct) بنسبة قليلة في فئتي المشاعر الإيجابية والمحايدة، حيث أن هنا النسبة الأقل في محتوى المشاعر الإيجابية في نموذج (GPT-4o mini) وزيادتها في محتوى المشاعر المحايدة تعد أفضل لأنها تدل على المزيد من التنوع في النموذج، ولكن الغياب شبه التام للمحتوى ذو المشاعر السلبية جعل أداء نموذج (LLAMA 3.1-70B Instruct) أفضل لتنوع النتائج على الفئات الثلاث من المشاعر فيه وتقارب النسب بينه وبين نموذج (GPT-4o mini) في الفئتين الإيجابية والمحايدة. أما نموذج (Gemini 1.5 F) كان له الأداء الأضعف من حيث تنوع المحتوى لطغيان المحتوى الإيجابي على السلبي والمحايد على حد سواء.

بعد النظر إلى ما ورد في العينة وتحليلها من حيث البنى النحوية التركيبية (الجملة الابتدائية)، يُستنتج الآتي:

- **النموذج الأول (GPT-4o mini):**

1. حقق نتائج متوسطة في توليد محتوى يبدأ ببنى جمل فعلية، (20%-56%) عبر المجموعات، (47%) للعينة كاملة.

2. حقق نتائج قليلة في توليد محتوى يبدأ ببنى جمل اسمية بالمقارنة مع أداؤه في توليد محتوى يبدأ ببنى جمل فعلية، (12%-28%) عبر المجموعات، (22%) للعينة كاملة.

3. أداء توليد المحتوى ذو بنى الجمل الاسمية متسق نوعاً ما عبر جميع المجموعات الفرعية عدا المجموعة الثانية.

4. حقق نتائج قليلة جداً في توليد محتوى يبدأ ببنى جمل ظرفية، (4%-24%) عبر المجموعات، (12%) للعينة كاملة.

5. أداء توليد المحتوى ذو بنى الجمل الظرفية غير متسق عبر المجموعات الفرعية.

6. النموذج الوحيد الذي كان فيه محتوى يبدأ بجملة شرطية، وولد مخرجاً واحداً فقط عبر المجموعات الفرعية الأربعة.

7. حقق النموذج نتائج بنسب تتراوح بين (8%-16%) في توليد محتوى جمل تبدأ بجار ومجرور، مما يشير إلى أداء أضعف في هذه الفئة مقارنة بالحالات الأخرى ولكنها نسبة معقولة بالنسبة لعدد الحالات المتوافرة لو كانت الفرص (النسب) متساوية بينها.

8. النموذج الوحيد الذي كان فيه محتوى يبدأ بجملة فعلية مسبوقة بحرف، نتائج بنسب تتراوح بين (0%–8%) عبر المجموعات، (4%) للعينة كاملة.

9. محتوى الجمل الفعلية المسبوقة بحرف غائبة تماماً في محتوى المجموعات (1, 2).

- **النموذج الثاني (Gemini 1.5 F):**

1. حقق نتائج متوسطة تميل للضعف في توليد محتوى يبدأ ببنى جمل فعلية، (4%–44%) عبر المجموعات، (23%) للعينة كاملة.

2. أداء توليد المحتوى ذو بنى الجمل الفعلية غير متسق عبر المجموعات الفرعية.

3. حقق نتائج عالية جداً في توليد محتوى يبدأ ببنى جمل اسمية بالمقارنة مع أداؤه في توليد محتوى يبدأ ببنى جمل فعلية، (40%–88%) عبر المجموعات، (64%) للعينة كاملة.

4. الأداء العالي في توليد محتوى يبدأ ببنى جمل اسمية يعني ضعف التنوع في الاستجابات من الأنواع الأخرى.

5. أداء توليد المحتوى ذو بنى الجمل الاسمية غير متسق عبر جميع المجموعات الفرعية، حيث تتراوح النسب للضعف بين المجموعات.

6. حقق النموذج نتائج بنسب تتراوح بين (4%–24%) في توليد محتوى جمل تبدأ بجار ومجرور، مما يشير إلى أداء أضعف في هذه الفئة مقارنة بالحالات الأخرى.

7. محتوى الجمل الفعلية المسبوقة بحرف ومحتوى الجمل الشرطية ومحتوى الجمل الظرفية غائبة تماماً في محتوى هذا النموذج.

- **النموذج الثالث (LLAMA 3.1-70B Instruct):**

1. حقق نتائج عالية جداً في توليد محتوى يبدأ ببنى جمل فعلية، (68%-96%) عبر المجموعات، (80%) للعينة كاملة.

2. يهيمن توليد محتوى يبدأ ببنى جمل فعلية بوضوح على هذه الفئة، مما يشير إلى أن النموذج مُحسَّن جيداً لتوليد هياكل الجمل الفعلية.

3. الأداء العالي في توليد محتوى يبدأ ببنى جمل اسمية يعني ضعف التنوع في الاستجابات من الأنواع الأخرى.

4. حقق نتائج ضعيفة في توليد محتوى يبدأ ببنى جمل اسمية بالمقارنة مع أداؤه في توليد محتوى يبدأ ببنى جمل فعلية، (0%-28%) عبر المجموعات، (10%) للعينة كاملة.

5. أداء توليد المحتوى ذو بنى الجمل الاسمية غير متسق عبر جميع المجموعات الفرعية.

6. حقق النموذج نتائج بنسب تتراوح بين (4%-20%) في توليد محتوى جمل تبدأ بجار ومجرور، مما يشير إلى أداء أضعف في هذه الفئة مقارنة بالحالات الأخرى.

7. محتوى الجمل الفعلية المسبوقة بحرف غائبة تماماً في محتوى المجموعات (1, 2).

8. محتوى الجمل الفعلية المسبوقة بحرف ومحتوى الجمل الشرطية ومحتوى الجمل الظرفية غائبة تماماً في محتوى هذا النموذج.

معيار مرجعي:

الرقم	المعيار
1	كلما ازدادت النسبة في توليد الاستجابات من نوع واحد من البنى النحوية في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أسوأ في توليد محتوى متنوع البنى النحوية
2	كلما انخفضت النسبة في توليد الاستجابات من نوع واحد من البنى النحوية في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أسوأ في توليد محتوى متنوع البنى النحوية
3	عندما تتقارب النسب في توليد الاستجابات من أنواع البنى النحوية الموجودة في المجموعات الفرعية/الكاملة كلما كان أداء النموذج أفضل في توليد محتوى متنوع البنى النحوية

جدول (9) معيار مرجعي من حيث تنوع البنى النحوية

(ملخص نتيجة 3):

يُستنتج أن نموذج (GPT-4o mini) تفوق في توليد محتوى متنوع من حيث البنية النحوية التركيبية في الجملة الابتدائية على النموذجين الآخرين، حيث يوجد فيه أنماط على عكس النموذجين الآخرين الذي تضمن المحتوى المولد منهما على ستة أنماط فقط عبر المجموعات الفرعية. وفيه أيضاً اتساق إلى حد ما عبر المجموعات الفرعية، على الرغم من تفوق فئة مجموعة الجمل الفعلية والاسمية على باقي المجموعات على التوالي. ويُستنتج أن نموذجي (LLAMA 3.1-70B Instruct) و(Gemini 1.5 F) متقاربين في الأداء من حيث توليد محتوى متنوع الموضوع، محتوى الجمل الاسمية واضحاً أكثر في (Gemini 1.5 F) من نموذج (LLAMA 3.1-70B Instruct) الذي يطغى عليه محتوى الجمل الفعلية بدرجة أكبر.

(ملخص نتيجة 4):

- يُستنتج أن (GPT-4o mini) هو النموذج الأفضل من حيث تنوع الموضوع وتنوع البنى النحوية التركيبية (الابتدائية) وعلى الرغم من أنه لم يكن في المقدمة من حيث توليد محتوى متنوع المشاعر لكن كان أداؤه جيد من توليد المحتوى الإيجابي والمحايد لكن يجب التنويه على الحاجة لتدريبه على المزيد من البيانات التي تحتوي على المشاعر السلبية والمحايدة (والسلبية بدرجة أكبر على وجه الخصوص).
- (GPT-4o mini) هو النموذج الأكثر موثوقية وثباتاً، مما يجعله الخيار الأفضل للتطبيقات العامة التي تتطلب التغطية والتنوع واتساع المعرفة.
- يعد (Gemini 1.5 F) خياراً ثانوياً جيداً في بعض الأحيان في المهام التي تتطلب تنوع المواضيع أو تنوع البنى النحوية التركيبية.
- يعد (LLAMA 3.1-70B Instruct) خياراً ثانوياً جيداً في بعض الأحيان في المهام التي تتطلب تنوع المشاعر.

- **ما التحديات والقيود التي تواجه خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي؟**

<u>إجابة السؤال الثاني:</u>

واجهت جميع النماذج المدروسة تحديات ملحوظة يُذكر منها:

1. بنية اللغة العربية النحوية المعقدة: أدى الهيكل الغني للغة العربية في بعض الأحيان إلى أخطاء نحوية، وخاصة في السياقات الأقل شيوعاً. هذه التعقيدات تتمثل في ترتيب الكلمات بدرجة أكبر للتأكيد على عناصر معينة، والاتفاق بين مكونات الجملة، التنسيق باستخدام الحروف، والجمل النسبية والجمل الشرطية وعلامات الحذف حيث تحذف العناصر المفهومة مثل الضمائر أو الروابط.. الخ. كل منها يتطلب محاذاة دقيقة للعناصر النحوية التي يصعب على النماذج تحديدها وتوليدها بأسلوب صحيح وخاصة في حالة عدم التدريب المخصص على اللغة العربية.

2. الغموض السياقي: أحياناً ما تفسر الخوارزميات العبارات الغامضة بطريقة خاطئة، مما ينتج عنه مخرجات تفتقر إلى الفروق الدقيقة السياقية.

3. البيانات المحدودة والتكرار: على الرغم من الأداء العام الأفضل لـ (GPT-4o mini)، إلا أن جميع النماذج واجهت صعوبات في التعامل مع تنوع المفردات، فيُلحظ تكرار الأفعال نفسها والبنى نفسها (مثال: 64 جملة بدأت ب "إن" في نموذج (Gemini 1.5 F) و39 جملة بدأت ب "يعتبر" في نموذج (LLAMA 3.1-70B Instruct) بسبب الفجوات في بيانات التدريب الخاصة بالمجال.

- **كيف تعمل خوارزميات الذكاء الاصطناعي المختلفة في توليد المحتوى العربي، وأي خوارزمية هي الأكثر فعالية؟**

<u>إجابة السؤال الثالث:</u>

تستفيد النماذج المدروسة جميعها من هياكل نموذج اللغة الكبيرة (LLM) مع الشبكات العصبية القائمة على المحولات. ومع ذلك، فإن أداءها كان متبايناً بسبب الاختلافات في حجم بيانات التدريب والجودة وطرائق الضبط الدقيق:

1. (GPT-4o mini): تفوق بسبب تعرضه لبيانات تدريب متعددة اللغات على نطاق واسع، وضبط دقيق قوي للفروق اللغوية الدقيقة، وتعامل أفضل مع العلاقات في الجمل العربية.

2. (Gemini 1.5 F): أظهر أداءً مقبولاً في المحتوى العام ولكنه تأخر في توليد محتوىً جيد من حيث تنوع المشاعر وواجه صعوبة في توليد بنى نحوية معقدة.

3. (LLAMA 3.1-70B Instruct): أظهر أداءً مقبولاً في المحتوى العام ولكنه تأخر في توليد محتوىً جيد من حيث تنوع المواضيع وواجه صعوبة في توليد بنى نحوية معقدة.

4. برز (GPT-4o mini) باعتباره الخوارزمية الأكثر فعالية لتوليد المحتوى العربي، متفوقة في التماسك والقواعد والسيولة الأسلوبية عبر مواضيع مختلفة.

- **ما تأثير حجم بيانات التدريب وجودتها في أداء خوارزميات الذكاء الاصطناعي في توليد المحتوى العربي؟**

<u>إجابة السؤال الرابع:</u>

أبرزت الدراسة وجود علاقة مباشرة بين حجم وجودة بيانات التدريب وأداء نماذج الذكاء الاصطناعي:

1.(GPT-4o mini): مكنتها مجموعة بيانات التدريب المتفوقة، والتي تضمنت مجموعة واسعة من النصوص العربية عالية الجودة (تشمل الأدب والمقالات الإخبارية والكتابة الفنية)، من توليد محتوى أكثر ثراءً ودقة.

2. (LLAMA 3.1-70B Instruct & Gemini 1.5 F): أعاقت مجموعات البيانات الأقل شمولاً هذه النماذج. مما أدى لتوليد محتوى منخفض التنوع عبر معايير الدراسة، مما أثر سلباً في الأداء، وخاصة في حالة (LLAMA 3.1-70B Instruct) الذي لا يدعم اللغة العربية.

4.3.1. مقارنة نتائج الدراسة الحالية مع الدراسات السابقة المشابهة:

1. إنها الدراسة الأولى (حسب علم الباحثة) التي درست النماذج (GPT-4o Mini) و(Llama 3.1-70B Instruct) و(Gemini 1.5 Flash) وقارنت بينها. وإنها الأولى في بحثها التطبيقي (مقارنة عمل نماذج لغوية) التي اعتمدت اللغة العربية كلغة البحث الأساسية فإن الدراسات السابقة اعتمدت الإنكليزية كلغة البحث مما يصعب على الباحثين العرب الآخرين الاستفادة منها بالقدر الكافي. كما كانت إنها الدراسة الأولى التي قارنت هذه النماذج عبر المهام المختلفة الآتية: تحليل المشاعر والبنية النحوية وتباين المواضيع.

4. تشابهت الدراسة مع دراسة (Al-Thubaity et al., 2023) من حيث معيار تحليل المشاعر لكن اختلفت معها بالمعايير الأخرى والنماذج المدروسة.

5. تشابهت نتائج الدراسة مع دراسة (Al-Thubaity et al., 2023) من حيث استناجهما تفوق نتائج نموذج Chatgpt في معيار تحليل المشاعر على النماذج المقارنة الأخرى، لكن اختلفت معها من حيث النماذج المقارنة الأخرى بين الدراستين واختلاف نسخة GPT بين الدراستين.

6. تشابهت نتائج الدراسة مع دراسة (El-Shangiti et al., 2024) من حيث استناجهما تفوق نتائج نموذج (Chatgpt) في إنتاج محتوى عالي التنوع، لكن اختلفت معها من حيث أن دراسة (El-Shangiti et al., 2024) تقارن بين نسخة GPT مضبوطة على بيانات في اللغة العربية ونسخة Turbo القياسية بينما دراستنا تقارن بين (Chatgpt) ونماذج أخرى. مما يوجه بأن سلسلة GPT عامة تعطي أداء مقبولاً بتوليد محتوىً عربياً جيداً ويمكن تحسينه بضبط النموذج على بيانات باللغة العربية لأداء أفضل.

7. اختلفت استراتيجية هذه الدراسة مع الدراسات السابقة بأنها ركزت أكثر على تقييم أداء النماذج أكثر مما اهتمت به الدراسات الأخرى في الدخول إلى عمليات بناء النموذج وضبطه بينما هذا البحث سعى لتقييم أداء النموذج أولاً لهدف تحديد نوع التقييم الذي يحتاجه النموذج وتوجهه لدراسات مستقبلية محتملة.

التوصيات والمقترحات:

توصي الباحثة بعد الدراسة السابقة بما يأتي:

- يجب أن يعتمد الذكاء الاصطناعي على أدوات وتقنيات لغوية متقدمة، مثل التدريب المسبق على مجموعات بيانات عربية متنوعة، وتحليل التبعيات لتحديد العلاقات النحوية، وآليات الانتباه لالتقاط التبعيات بعيدة المدى.
- تقديم مجموعات بيانات خاصة بتوليد محتوى اللغة العربية في أثناء عمليات الضبط الدقيق لمعالجة نقاط الضعف في النماذج وتحسين قدرتها على التعامل مع اللغة العربية وأساليبها.
- تطوير بروتوكولات تدريب محددة لمعالجة التحديات المتأصلة في الصرف العربي، والنظر في دمج القواعد اللغوية وأدوات التحليل النحوي في تدريب الذكاء الاصطناعي لتعزيز الدقة النحوية.
- يمكن أن تساعد المعالجة اللاحقة القائمة على القواعد في التحقق من صحة الجمل التي أنشئت بوساطة الذكاء الاصطناعي وتحسينها، مما يضمن توافقها مع قواعد النحو العربية الغنية والمعقدة.
- تدريب النماذج على نصوص عربية أطول وأكثر تعقيداً لتعزيز فهمها للعلاقات عبر جمل أو فقرات متعددة.

- تحديث مجموعات بيانات التدريب باستمرار لتشمل المحتوى الحالي عالي الجودة باللغة العربية، مما يضمن بقاء النماذج ذات صلة ودقة بمرور الوقت.
- التأكد من تحديد أي تحيزات في بيانات التدريب والتخفيف منها لإنتاج محتوى محايد وشامل.
- دمج الملاحظات النوعية من المتحدثين الأصليين باللغة العربية لتقييم أفضل لطبيعية المحتوى الناتج وملاءمته الثقافية.
- التعاون مع اللغويين واللغويين الحاسوبيين وباحثي الذكاء الاصطناعي المتخصصين في اللغة العربية لمواءمة أفضل للنماذج مع الفروق الدقيقة في قواعد اللغة العربية والنحو والسياق الثقافي.
- تشجيع المؤسسات الحكومية والخاصة على الاستثمار في تطوير نماذج الذكاء الاصطناعي المدعومة والمخصصة للغة العربية.

الخاتمة:

وفي الختام، سلطت دراسة " تأثير خوارزميات الذكاء الاصطناعي (LLM) المدربة في توليد المحتوى باللغة العربية" الضوء على التطورات والتحديات المهمة التي تفرضها نماذج اللغة الكبيرة في إنشاء المحتوى العربي. وأظهرت مقارنة نماذج الذكاء الاصطناعي المختلفة أن نماذج اللغة الكبيرة تُظهر قدرة قوية على توليد نص عربي متنوع عبر مواضيع مختلفة. ومع ذلك، لا تزال التحديات قائمة في مجالات مثل الاختلافات اللهجية، والفروق الثقافية، والحفاظ على الدقة اللغوية، وخاصة عند التعامل مع أشكال معقدة أو أقل توحيداً للغة العربية.

قدمت هذه الدراسة فرصة رائعة بمقارنة شاملة لنقاط القوة والضعف في كل نموذج، ووفرت فهماً واضحاً للنموذج الذي يعمل بطريقة أفضل في سيناريوهات

مختلفة، ونتائج قابلة للقياس يمكن استخدامها لاتخاذ قرارات مستنيرة حول استخدام نماذج اللغة الكبيرة لمهام توليد المحتوى باللغة العربية والحاجة إلى مواصلة البحث في تقاطع الذكاء الاصطناعي واللغة والثقافة.

وعلاوة على ذلك، في حين أن نماذج اللغة الكبيرة قادرة على إنتاج محتوى عالي الجودة والتنوع، فإن هذه الدراسة أكدت على الحاجة إلى التحسين المستمر وتكييف نماذج الذكاء الاصطناعي لتعامل أفضل مع تعقيدات اللغة العربية، بما في ذلك النحو والصرف. وتشير النتائج إلى أنه في حين تُظهر خوارزميات الذكاء الاصطناعي وعداً كبيراً لتوليد المحتوى باللغة العربية، لا تزال هناك فجوة في مطابقة الفهم والإبداع على المستوى البشري. ستكون التحسينات المستقبلية في تنوع البيانات وتقنيات تدريب النماذج ومقاييس التقييم حاسمة في تعزيز فعالية وقابلية تطبيق المحتوى العربي الذي ينشئ بوساطة الذكاء الاصطناعي.

في نهاية المطاف، قدمت هذه الدراسة رؤى قيمة حول القدرات الحالية إنشاء المحتوى العربي وتقترح حلول أكثر دقة تعتمد على الذكاء الاصطناعي للمجتمعات الناطقة باللغة العربية.

المراجع:

المراجع الأجنبية:

- (2024). ***The rise of Large Language Models: from fundamentals to application.*** Management Solutions.
- Abdul-Mageed et al., (2021). ***ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic.*** Canada: The University of British Columbia, Natural Language Processing Lab.
- Agarwal et al.. (2024). ***Copilot Evaluation Harness: Evaluating LLM-Guided Software Programming.*** USA: Microsoft. p1.
- Al-Ajlan, A.; Al-Khalifa, H.; & Al-Salman, A.. (2008). ***Towards the Development of an Automatic Readability Measurements for the Arabic Language.*** Proceedings of the 3rd International Conference on Digital Media.
- Al-Thubaity et al.. (2023). ***Evaluating ChatGPT and Bard AI on Arabic Sentiment Analysis***. Saudi Arabia: King Abdulaziz City for Science and Technology & Imam Mohammad ibn Saud University & Qassim University.
- Alpaydin, E. (2014). ***Introduction to Machine Learning.*** USA: MIT Press.
- Antoun et al.. (2020). ***AraBERT: Transformer-based Model for Arabic Language Understanding***. Lebanon: American University of Beirut.
- Antoun et al.. (2021). ***ARAGPT2: Pre-Trained Transformer for Arabic Language Generation.*** Lebanon: American University of Beirut.
- Antoun et al.. (2024). ***From Text to Source: Results in Detecting Large Language Model-Generated Content***. France, Paris: Inria.
- Aryan et al.. (2024). ***The Costly Dilemma: Generalization, Evaluation and Cost-Optimal Deployment of Large Language Models.***
- Ba et al.. (2016). ***Layer Normalization***. Canada: University of Toronto.
- Barone et al.. (2017). ***Regularization techniques for fine-tuning in neural machine translation.*** Scotland: The University of Edinburgh, School of Informatics. p1.

- Behera, Santosh Kumar; Nayak, Mitali M. (2020). ***Natural Language Processing for Text and Speech Processing: A Review Paper.*** India: Siksha 'O' Anusandhan, Department of Computer Science and Engineering. p2-3.
- Bender et al.. (2021). ***On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?***. Proceedings of the 2021 ACM FAccT Conference.
- Bengio, Y. et al.. (2009). ***Proceedings of the 26th International Conference on Machine Learning.*** Canada: Association for Computing Machinery.
- Bilgihan, A.; Kandampully, J.; & Zhang, T. (2016). ***Towards a unified customer experience in online shopping environments: Antecedents and outcomes***. International Journal of Quality and Service Sciences.
- Bilski, Adrian. (2011). ***A Review of Artificial Intelligence Algorithms in Document Classification***. Poland: International Journal of Electronics and Telecommunications.
- Boulki et al.. (2024). ***Evaluating the Effectiveness of Pre-trained Language Models in Predicting the Helpfulness of Online Product Reviews.*** Netherlands: Tilburg University.
- Brown et al.. (2020). ***Language Models are Few-Shot Learners***. USA: OpenAI.
- Cheung, Ming. (2024). ***A Reality check of the benefits of LLM in business.*** China: The Lane Crawford Joyce Group, Beta Labs.
- Chockalingam et al.. (2023). ***A Beginner's Guide to Large Language Models - Part 1***. Nividia.
- Chung et al.. (2014). ***Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling***. Canada: University of Montreal.
- Davenport et al.. (2020). ***How artificial intelligence will change the future of marketing.*** Journal of the Academy of Marketing Science.
- De Lange et al.. (2021). ***A continual learning survey: Defying forgetting in classification tasks***. IEEE.

- Deng, Jianyang; Lin, Yijia. (2022). ***The Benefits and Challenges of ChatGPT: An Overview.*** China: Bloco China Sociedade Unipessoal Limitada & City University of Macau, Faculty of Finance. p82.
- Devlin et al.. (2018). ***BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.*** USA: Google AI Language.
- Du et al.. (2024). ***Reinforcement Learning with Large Language Models (LLMs) Interaction For Network Services.*** IEEE.
- Duan et al.. (2021). ***A Survey of Few-Shot Learning: An Effective Method for Intrusion Detection.*** China: Beijing Information Science and Technology University (Beijing Advanced Innovation Center for Materials Genome Engineering & Laboratory of Data Science and Information Studies) & Beijing Laboratory of National Economic Security Early-Warning Engineering.
- El-Shangiti et al.. (2024). ***Arabic Automatic Story Generation with Large Language Models***. Canada: The University of British Columbia & KSA: Mohamed bin Zayed University of Artificial Intelligence & Invertible AI.
- Farghaly; Shaalan, K.. (2009). ***Arabic natural language processing: Challenges and solutions***. ACM Transactions on Asian Language Information Processing.
- Finn, C.; Abbeel, P.; Levine, S. (2017). ***Model-agnostic meta-learning for fast adaptation of deep networks***. In Proceedings of the 34th International Conference on Machine Learning. Australia, Sydney: ICML & USA: University of California & OpenAI.
- Fleshman, William; Van Durme, Benjamin. (2023). ***Toucan: Token-Aware Character Level Language Modeling.*** USA: John Hopkins University.
- Fred, Williams. (2023). ***Societal Impact and LLMAdoption: A Comprehensive Examination***. OSF.
- Geis et al.. (2019). ***Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement***. Canada: Canadian Association of Radiologists Journal.

- Goldberg, Y. (2017). ***Neural network methods for natural language processing***. Synthesis Lectures on Human Language Technologies.
- Goodfellow et al.. (2014). ***Generative adversarial networks. In Advances in Neural Information Processing Systems.*** Google Brain & University of Montreal.
- Guo, Yihao. (2024). ***Instruction Fine-Tuning: The Key to Professional, HighQuality Automated Writing.*** China, Macao: Macau University of Science and Technology, The Faculty of Innovation Engineering.
- Habash, N.. (2007). ***Arabic morphological representations for machine translation***. Columbia University, Center for Computational Learning Systems.
- Havrilla et al.. (2024). ***Teaching Large Language Models to Reason with Reinforcement Learning.*** USA: Meta & Georgia Institute of Technology & StabilityAI & UC Berkeley.
- He et al.. (2021). ***On the Effectiveness of Adapter-based Tuning for Pretrained Language Model Adaptation.*** Singapore: DAMO Academy & Alibaba Group & Nanyang Technological University & National University of Singapore & Singapore University of Technology and Design. p1.
- Hinton, G.; & Sejnowski, T. J. **(1999)**. ***Unsupervised Learning.*** USA: MIT Press.
- Hinton, Geoffrey; Vinyals, Oriol; Dean, Jeff. (2015). ***Distilling the Knowledge in a Neural Network***. USA: Google.
- Hochreiter, Sepp; Schmidhuber, Jurgen. (1997). ***Long Short-Term Mempry***. USA: Massachusetts Institute of Technology.
- Jaiswal, A. et al.. (2021). ***IEEE Transactions on Neural Networks and Learning Systems.*** IEEE.
- Jiao et al.. (2024***). Navigating LLM Ethics: Advancements, Challenges, and Future Directions.*** USA, Texas, Austin: The University of Texas at Austin, The School of Architecture, Urban Information Lab. p9,13.

- Johri et al.. (2020). ***Natural Language Processing: History, Evolution, Application and Future Work.*** India: Galgotias University, & Jordan: Yarmouk University & Uzbekistan: Samarkand State University. p2-4.
- Jospin et al.. (2022). ***Hands-on Bayesian Neural Networks – A Tutorial for Deep Learning Users.*** Australia: University of Western Australia & Murdoch University & Monash University. p1.
- Kaddour et al.. (2023). ***Challenges and Applications of Large Language Models.*** University College London & University of Cambridge & UK Health Security Agency & EleutherAI & Stability AI & Meta AI Research & InstaDeep.
- Karoo; Krishna, Jogi; Meghraj. (2023). ***Syntactic and Semantic Analysis in Natural Language Processing: Unveiling the Underlying Mechanisms.*** India: Gondwana University, Department of Computer Science. p1.
- Khalid et al.. (2024). ***The Role of Language Models in Modern Healthcare: A Comprehensive Review***. USA: HRISTUS Santa Rosa Hospital New Braunfels & Pakistan: Riphah International University.
- Kingma, D. P., & Welling, M. (2014). ***Auto-encoding variational bayes.*** In Proceedings of the 2nd International Conference on Learning Representations. Netherlands: University van Amsterdam, Machine Learning Group.
- Kirkpatrick et al.. (2017). ***Overcoming catastrophic forgetting in neural networks***. United Kingdom, London: Imperial College London, Bioengineering department & DeepMind.
- Koubaa et al.. (2024). ***ArabianGPT: Native Arabic GPT-Based Large Language Model***. KSA, Riyadh: Prince Sultan University, Robotics and Internet-of-Things Lab.
- Kudo, Taku; Richardson, John. (2018). ***SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing.*** USA: Google.
- Larkey, L. S.; Ballesteros, L.; & Connell, M. E. (2002). ***Improving stemming for Arabic information retrieval: Light stemming and co-occurrence analysis***.

- Lee et al.. (2019). ***BioBERT: a pre-trained biomedical language representation model for biomedical text mining***.
- Lewis et al.. (2019). ***BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension.*** Facebook AI.
- Liddy, Elizabeth. (2001). ***Natural Language Processing***. USA: Syracuse University, School of Information Study. p1, 5-7, 8, 10 11-12.
- Li et al.. (2021). ***Pretrained Language Models for Text Generation: A Survey.*** China.
- Li et al.. (2024). ***Infusing Human Feedback into Intermediate Prompting Steps of Large Language Models.*** USA: Clemson University. p3.
- Ling et al.. (2024). ***Domain Specialization as the Key to Make Large Language Models Disruptive: A Comprehensive Survey.*** USA: Emory University & Rutgers University & University of Pittsburgh & George Mason University & University of California & Duke University & NEC Labs America & Blackrock, Inc. & Microsoft. p2.
- Liu Y et al.. (2019). ***RoBERTa: A Robustly Optimized BERT Pretraining Approach.*** USA, WA, Seattle: University of Washington Paul G. Allen School of Computer Science & Engineering.
- Liu, Y; et al.. (2024). ***Understanding LLMs: A Comprehensive Overview from Training to Inference.*** Netherlands: Elsevier.
- Liu, Y. et al.. (2024). ***Datasets for Large Language Models: A Comprehensive Survey.*** China: South China University of Technology & INTSIG Information Co., Ltd & INTSIG-SCUT Joint Lab on Document Analysis and Recognition. p5.
- Marie-Sainte et al.. (2017). ***Arabic Natural Language Processing and Machine Learning-based Systems***. KSA: Prince Sultan University, College of Computer and Information Sciences, Computer Science Department & King Saud University, College of Computer and Information Sciences, Information Technology Department.

- McCarthy, J. (1959). ***Programs with common sense: In Proceedings of the Teddington Conference on the Mechanization of Thought Processes.***
- Meduri, Sudeep. (2024). ***Revolutionizing Customer Service: The Impact of Large Language Models on Chatbot Performance.*** India: Indian Institute of Technology.
- Mohammad, Atif Farid; Clark, Bryan; Hegde, Ramya. (2023). ***Large Language Model (LLM) & GPT, A Monolithic Study in Generative AI.*** USA: University of North Carolina.
- Mousi et al.. (2024). ***ARADICE: Benchmarks for Dialectal and Cultural Capabilities in LLMs.*** Canada: University of New Brunswick & Qatar: Qatar Computing Research Institute & UAE: American University of Sharjah.
- Nagoudi et al.. (2022). ***AraT5: Text-to-Text Transformers for Arabic Language Generation.*** Canada: The University of British Columbia & Deep Learning and Natural Language Processing Group.
- Nagoudi et al.. (2023). ***JASMINE: Arabic GPT Models for Few-Shot Learning.*** Canada: The University of British Columbia, Deep Learning – Natural Language Processing Group & UAE: MBZUAI Department of Natural Language Processing – Department of Machine Learning.
- Naveed et al.. (2024). ***A Comprehensive Overview of Large Language Models***. Elsevier.
- Pan, Sinno Jialin; Yang, Qiang. (2010). ***A Survey on Transfer Learning***. IEEE.
- Parthasarathy et al.. (2024). ***The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Exhaustive Review of Technologies, Research, Best Practices, Applied Research Challenges and Opportunities.*** Ireland, Dublin: University College Dublin, CeADAR, Ireland's Centre for AI. p9.
- Pires et al.. (2023). ***One Wide Feedforward is All You Need***. Equall & Apple.

- Potter, Yujin; Potter, Michael; Song, Dawn. (2024). ***As an AI, I believe AI models should be open source***. USA: UC Berkeley, Department of Electrical Engineering and Computer Sciences
- Prechelt, Lutz. (2002). ***Early Stopping - But When?.*** Germany: Karlsruhe Institute of Technology.
- Raffel et al.. (2019). ***Exploring the limits of transfer learning with a unified textto-text transformer.*** USA: Google.
- Raiaan et al.. (2024). ***A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges.*** Bangladesh: United International University - Bangladesh University of Engineering and Technology & Finland: Lahti University of Technology & Australia: Charles Darwin University.
- Reizinger et al.. (2024). ***Position: Understanding LLMs Requires More Than Statistical Generalization.*** Germany: Max Planck Institute for Intelligent Systems - ELLIS Institute - Tübingen AI Center & UK: University of Cambridge & Switzerland: Department of Computer Science ETH Zurich - Max Planck ETH Center for Learning Systems.
- Russell, Stuart; Norvig, Peter. (1995). ***Artificial Intelligence: A Modern Approach***. USA.
- Sanh et al.. (2020). ***DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter***. Hugging Face.
- Sazli, Murat H. (2006). ***A Brief Review of Feed Forward Neural Networks.*** Turkey: Ankara University, Faculty of Engineering, Department of Electronics Engineering.
- Schuster, Mike; Nakajima, Kaisuke. (2012). ***Japanese and Korean Voice Search.*** USA: Google.
- Sener, Ozan; Koltun, Vladlen. (2019). ***Multi-Task Learning as Multi-Objective Optimization.*** USA: Intel Lab.
- Sennrich et al.. (2015). ***Neural Machine Translation of Rare Words with Subword Units.*** Scotland: University of Edinburgh, School of Informatics.

- Settles, B. (2009). ***Active Learning Literature Survey.*** USA: University of Wisconsin-Madison.
- Shaalan, K.. (2005). ***An intelligent computer-assisted language learning system for Arabic learners***.
- Sheikh, Haroon; Schrijvers, Erik; Prins, Corien. (2023). ***Mission AI.*** Netherlands: The Hague, the Netherlands Scientifc Council for Government Policy.
- Sun et al.. (2020). ***MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices.*** USA: Carnegie Mellon University & Google Brain. p1.
- Turney, Peter D. (2005). ***Measuring Semantic Similarity by Latent Relational Analysis.*** Canada: National Research Council Canada, Institute for Information Technology. p1-2.
- Vaswani et al.. (2017). ***Attention Is All You Need.*** USA: Google.
- Vavekanand, Raja; Kumar, Sanjay. (2024). ***LLMEra: Impact of Large Language Models***. Pakistan: University of Sindh, Datalink Research and Technology Lab.
- Wachsmuth, Henning. (2023). ***Statistical Natural Language Processing***. Germany: University of Hannover. p7-9.
- Wang et al.. (2024). ***Understanding User Experience in Large Language Model Interactions.*** China: Tsinghua University & Canada: University of Montreal. p1.
- White et al.. (2023). ***A prompt pattern catalog to enhance prompt engineering.*** Shanghai AI Laboratory & Jilin University & Zhejiang University & The Chinese University of Hong Kong.
- Wu et al.. (2023). ***Multimodal Large Language Models: A Survey.*** China: Jinan University & Pazhou Lab & South China University of Technology & USA: The University of Illinois Chicago. p1.

- Xhonneux et al.. (2024). ***Efficient Adversarial Training in LLMs with Continuous Attacks.*** Canada: University of Montréal & Germany: Technical University of Munich & Microsoft Research. p2-3.
- Xu et al.. (2024). ***A Survey on Multilingual Large Language Models: Corpora, Alignment, and Bias.*** The Fundamental Research Funds for the Central Universities. p1.
- Yang et al.. (2019). ***XLNet: Generalized Autoregressive Pretraining for Language Understanding.*** USA: Carnegie Mellon University & Google AI Brain Team.
- Yang et al.. (2020). ***FinBERT: A Pretrained Language Model for Financial Communications.*** China: Hong Kong University of Science and Technology, School of Business and Management.
- Yu et al.. (2024). ***What Makes a High-Quality Training Dataset for Large Language Models: A Practitioners' Perspective.*** China: Wuhan University of Technology & Zhejiang University & Huawei & Canada: University of Ottawa & Australia: Monash University. p5.
- Zhang et al.. (2016). ***LibN3L: A Lightweight Package for Neural NLP.*** China: Heilongjiang University & Singapore: Singapore University of Technology and Design. p1.
- Zhang et al.. (2024). ***LA-UCL: LLM-Augmented Unsupervised Contrastive Learning Framework for Few-Shot Text Classification.*** China: Tianjin University, College of Intelligence and Computing. p8.
- Zhao et al.. (2023). ***A Survey of Large Language Models.***
- Zhou, Zhi-Hua. (2012). ***Ensemble Methods: Foundations and Algorithms***. England: Chapman & Hall.
- Zhou et al.. (2024). ***A Survey on Efficient Inference for Large Language Models.*** IEEE. p1.
- Zou, Kai; Wei, Jason. (2019). ***EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks.*** USA: Georgetown University & Dartmouth University.

- Amazon. (2023). https://aws.amazon.com/what-is/recurrent-neural-network/#:~:text=A%20recurrent%20neural%20network%20(RNN,complex%20semantics%20and%20syntax%20rules.
- Babenko, Konstantin. (2024). https://www.linkedin.com/pulse/choosing-right-llm-guide-proprietary-vs-open-source-babenko-ph-d--gcese
- Chatgptguide. (2024). https://www.chatgptguide.ai/2024/02/29/what-is-inference-llms-explained/
- ContextAi. https://context.ai/compare/llama3-1-70b-instruct-v1/gemini-flash
- Elastic. (2024). https://www.elastic.co/what-is/large-language-models
- Google Deepmind. (2024). https://deepmind.google/technologies/gemini/
- Hazelcast. (2024). https://hazelcast.com/foundations/ai-machine-learning/machine-learning-inference/
- Hugging Face. (2024). https://huggingface.co/meta-llama/Llama-3.1-70B-Instruct
- Linkedin. (2024). https://www.linkedin.com/advice/3/how-can-you-develop-intelligent-chatbot-trebf
- Open AI. (2022). https://openai.com/index/whisper/
- Open AI. (2024). https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/

المراجع العربية:

- استيتية، شريف سمير. (2008). *اللسانيات، المجال، والوظيفة، والمنهج*. الأردن، إربد: عالم الكتب الحديث. ص527.
- برينيس; ربيع، بشار; إبراهيم. (2022). *دور اللسانيات الحاسوبية في خدمة اللغة العربية*. الجزائر: جامعة بسكرة، مخبر اللسانيات واللغة العربية. ص 8-9.
- بشار، إبراهيم. (2020). *اللسانيات الحاسوبية التأسيس الغربي والتلقي العربي*. الجزائر: جامعة محمد خيضر، قسم الآداب واللغة العربية.
- بعلبكي، رمزي منير. (1990). *معجم المصطلحات اللغوية*. لبنان، بيروت: دار العلم للملايين.
- بلحداد; ايمان، شتوح; زهور. (2021). *الجهود اللسانية الحاسوبية العربية: قراءة في المنهج وآفاق البحث*. الجزائر: جامعة باتنة.

- داود، محمد محمد. (2001). *العربية وعلم الحديث*. مصر، القاهرة: دار غريب، القاهرة.
- دويدري، رجاء وحيد. (2000). *البحث العلمي أساسياته النظيرة وممارسنه العملية*. سوريا: دار الفكر العربي. ص183.
- السباعي; ريم. (2022). *دور اللسانيات الحاسوبية في مجال تعليم اللغة العربية*. سوريا: الجامعة الافتراضية السورية. ص19-24.
- المقدسي، مطهر بن طاهر. (1959). *البدء والتاريخ*. مصر، القاهرة: دار الفكر العربى للطباعة والنشر.
- مهديوي، عمر. (2008). *توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية مقاربة لسانية حاسوبية*. المغرب، الدار البيضاء: جامعة الحسن الثاني، كلية الآداب والعلوم الإنسانية، شعبة اللغة العربية وآدابها، وحدة علوم اللغة العربية والمعجميات.
- اليوم السابع. (2024). 10توصيات لأول مؤتمر عن الذكاء الاصطناعي وانتهاك الملكية الفكرية لوزارة العدل
- مركز الاتحاد للأخبار. (2024). المؤتمر السنوي الدولي الرابع لـ«ترندز» يبحث الأمن المستدام والذكاء الاصطناعي

Abstract:

Study Title: The Impact of Trained Artificial Intelligence Algorithms (LLM) on Content Generaion in Arabic Language.

Study Objective: To investigate the impact of trained AI algorithms on Content Generaion in Arabic Language, focusing on evaluating their performance, identifying challenges and limitations, and exploring potential applications. Specifically, this study aims to evaluate the performance of three modern AI models in generating accurate Arabic content, and generating contextually correct relevant Arabic text. This study examines the effectiveness of AI algorithms in generating diverse, high-quality Arabic content.

Study Methodology: The study adopted a descriptive-analytical approach, which enabled a comprehensive evaluation of the linguistic models (LLAMA, Gemini, and GPT4). Their performance was evaluated by generating a dataset of Arabic text and comparing them across Arabic language tasks (content diversity, sentiment analysis, and syntactic diversity). The results of the experiment were analyzed using statistical methods.

Study Results: The study demonstrates the effectiveness of AI algorithms in generating Arabic content:

1. (LLAMA 3.1-70B Instruct) excelled in sentiment diversity, thanks to balanced ratios of positive, negative, and neutral sentiments. (GPT-4o mini) performed best in both positive and neutral sentiments, but

lacked negative sentiments. (Gemini 1.5 F) underperformed due to the dominance of positive content.

2. (GPT-4o mini) clearly outperformed, while the performance of (LLAMA 3.1-70B Instruct) and (Gemini 1.5 F) was close.
3. (GPT-4o mini) excelled in providing diverse and consistent patterns, while the other models lacked diversity and relied on limited patterns.
4. (GPT-4o mini) is the best model overall, but needs improvement for negative sentiments. (LLAMA 3.1-70B Instruct) is a good choice for diverse sentiments. (Gemini 1.5 F) is a good secondary choice for diverse topics and grammatical structures.