معالجة تعقيدات اللغة العربية باستخدام تقنيات الذكاء الاصطناعي: التحديات والحلول المستقبلية

صابرين علي حسين[1]، اسيل حمود حمزة[2*] ، فرح علاء عبد الحسن [3]
جامعة بابل، كلية التربية الأساسية[1]، جامعة بابل، كلية القانون[2]، جامعة بابل، كلية تكنلوجيا المعلومات (العراق) [3]

Addressing the Complexities of the Arabic Language with AI Techniques: Challenges and Future Solutions

Sabreen Ali Hussein[1], Aseel Hamoud Hamza[2*], Farah A. Hasan [3]
https://orcid.org/0000-0003-4338-4078[1] , https://orcid.org/0000-0003-2124-2108[2], https://orcid.org/0009-0008-5082-5947[3]
University of Babylon, College of Basic Education[1], University of Babylon, College of Law [2],University of Babylon, Computer science College of IT /Iraq)[3]
qur.sabreen.alshammary@uobabylon.edu.iq[1], aseel.hamod@ uobabylon.edu.iq[2], inf883.frh.alaa@ uobabylon.edu.iq[3]

تاريخ الاستلام: 17 / 1 /2025 تاريخ القبول: 02 / 03 / 2025 تاريخ النشر: 01 / 06 / 2025

الملخص:

تُعتبر اللغة العربية أكثر اللغات تعقيدًا من حيث البناء والتركيب، حيث تتميز بخصائص متعددة مثل التصريف والاشتقاق، بالإضافة إلى تنوع اللهجات وتعدد المعاني. كل هذه العوامل تتطلب أنظمة ذكاء اصطناعي متطورة قادرة على استيعاب هذا الكم الهائل من التعقيدات. ونظرًا لأن معالجة اللغات الطبيعية تُعد من أبرز مجالات الذكاء الاصطناعي، فإنه من الضروري أن يشهد هذا المجال تطورًا يتماشى مع ما تتمتع به اللغة العربية من خصائص فريدة وتعقيدات.

في هذا البحث، سوف نسلط الضوء على تقنيات الذكاء الاصطناعي التي تساهم في تعزيز ودعم اللغة العربية، بالإضافة إلى تحليل التحديات الحالية التي تواجه هذه التقنيات. سنعمل على فهم كيفية معالجة خصائص اللغة العربية والخوارزميات المتاحة، بما يتماشى مع التطلعات المستقبلية.

كلمات مفتاحية: الذكاء الاصطناعي، اللغة العربية،. معالجة اللغات الطبيعية،. التحديات.

Abstract:

Arabic language is one of the most complex languages. in terms of structure and composition, It has inflection, derivation, diversity of dialects, and different typee of meaning . All of this requires sophisticated artificial intelligence systems that can understand this wide variety of complicity. Since one major type of artificial intelligence is natural language processing, natural language advancements in this field must be compatible with the content and art of the Arabic language. We focus in research on artificial intelligence technologies, techniques and tools that help in the promotion and support of the Arabic language analysis, and the problems faced by these technologies in understanding and dealing with the characteristics of the Arabic language. We will look at the current algorithms and those that meet the future needs.

Keywords: Artificial Intelligence; Arabic language; Natural Language Processing; AI technologies; Challenges.

مقدمة:

لقد ساهم الذكاء الاصطناعي في تحسين تحليل النصوص العربية من خلال إعادة تشكيل وفهم المحتوى وإنتاجه. حيث يُعتبر تعلم الآلة أداة تمكّن الآلات من التعلم من البيانات واتخاذ القرارات بأقل تدخل بشري. كما أن استخدام الذكاء الاصطناعي يُعزز من سرعة تحليل البيانات والتعرف على الأنماط واتخاذ القرارات المتعلقة بالمحتوى العربي، وتوفر الوقت وتقليل الحاجة إلى التدخل اليدوي.

ومع ذلك، هناك بعض التحديات مثل التحيز الخوارزمي (Algorithmic Bias) وغياب الشفافية في اتخاذ القرارات. ورغم هذه الصعوبات، يُعتبر الذكاء الاصطناعي واعدًا في تحسين جودة النصوص العربي. كما أن أدوات التدقيق والتحرير المدعومة بالذكاء الاصطناعي تُساهم في رفع جودة النصوص من خلال اكتشاف الأخطاء واقتراح التحسينات، على الرغم من أنها قد تنشر أحيانًا معلومات مضللة إذا لم تُراقب بعناية. لذا، من الضروري تحقيق التوازن بين الكفاءة والدقة لضمان جودة المحتوى العربي.

على الرغم من الإمكانات التي يوفرها الذكاء الاصطناعي لتوسيع السوق وإنشاء محتوى متعدد اللغات، بالإضافة إلى تقديم أدوات تحليل مبتكرة، تبقى تعقيدات اللغة العربية والاعتبارات الأخلاقية مثل التمييز وانتهاك الخصوصية من أكبر التحديات. يتطلب هذا المجال تحقيق توازن بين التقدم التكنولوجي والاعتبارات الأخلاقية لإنتاج محتوى عربي عالي الجودة [1], [9], [7], [11], [10], [3].

مشكلة البحث

تركز على استكشاف التحديات المرتبطة بخوارزميات الذكاء الاصطناعي المستخدمة في توليد النصوص العربية، مثل تعقيد اللغة العربية، ونقص البيانات، وانخفاض دقة النصوص، وانحياز النموذج، والمخاوف المتعلقة بالخصوصية، والتأثير على الثقة في المحتوى.

اهداف البحث

1. تحليل الإيجابيات والسلبيات لخوارزميات الذكاء الاصطناعي في توليد النصوص العربية.
2. تقديم توصيات لتحسين أداء الخوارزميات والتقليل من التحديات المرتبطة بها.
3. تحديد التحديات الأخلاقية والتقنية المرتبطة بهذه الخوارزميات.

اهمية البحث

تتركز الاهمية من الحاجة إلى استخدام تقنيات الذكاء في توليد النصوص العربية بشكل فعال وأمن. يمكن أن تساهم النتائج في تحسين التطبيقات الحالية وتطوير حلول جديدة مصممة لتلبية الاحتياجات الخاصة باللغة العربية.

منهجية البحث

تستند الدراسة إلى مراجعة البحوث المتعلقة بخوارزميات الذكاء الاصطناعي وتوليد النصوص، بالإضافة إلى تحليل البيانات والتجارب العملية لتحديد الإيجابيات والتحديات. يتم استخدام أساليب البحث النوعي والكمي للحصول على نتائج متكاملة.

المبحث الاول

استخدام الذكاء الاصطناعي في اللغة العربية

المطلب الاول :ماهية الذكاء الاصطناعي

1-الذكاء الاصطناعي[25]:

الذكاء هو القدرة على التعلم وفهم المفاهيم والتكيف وحل المشكلات. يتميز الأشخاص الأذكياء بقدرتهم على التفكير النقدي، التحليل، الابتكار، والقدرة على التعامل مع المعلومات الجديدة، فالذكاء سمة متعددة الأبعاد وتتجلى بطرق مختلفة في سياقات متنوعة.

2-أنواع الذكاء لاصطناعي[25]:

الذكاء اللفظي: امكانية استخدام اللغة بفعالية (المكتوبة أو المنطوقة).

الذكاء الرياضي والمنطقي: امكانية التفكير وحل المسائل الرياضية.

الذكاء البصري والمكاني: امكانية تصور وتحليل الأشكال والمساحات.

الذكاء الموسيقي: امكانية فهم الموسيقى وتأليفها.

الذكاء الجسدي الحركي: امكانية تحريك الجسم بمهارة.

الذكاء الشخصي: القدرة على فهم الذات والسيطرة على المشاعر.

الذكاء الاجتماعي: امكانية فهم الآخرين والتفاعل معهم.

الذكاء الطبيعي: امكانية التعرف على الأنواع والأنماط في العالم الطبيعي.

3-تطبيق الذكاء لاصطناعي [25]

1. التعليم: استخدام استراتيجيات تعليمية مناسبة لأنواع الذكاء المختلفة لتحسين تجربة التعلم.
2. الأعمال: تطبيق الذكاء الاصطناعي في اتخاذ القرارات وتحليل البيانات.
3. الرعاية الصحية: استخدام الذكاء في تحليل البيانات الطبية وتشخيص الأمراض.
4. تكنولوجيا المعلومات: تطوير البرمجيات والأنظمة الذكية التي تسهل العمليات اليومية.
5. الأمن السيبراني: استخدام الذكاء في حماية المعلومات ولتحليل التهديدات.

المطلب الثاني :توصية الأدبيات في استخدام الذكاء الاصطناعي

1-الدراسات السابقة

يتم استعراض استخدام الذكاء الاصطناعي ومعالجة اللغات الطبيعية في برامج الترجمة عبر الإنترنت، مع التركيز على اللغة العربية. ويناقش التحديات التي تواجهها برامج الترجمة هذه، مثل تعقيد اللغة العربية وتنوع لهجاتها وعدم وجود قاعدة بيانات شاملة للغة العربية. كما يناقش تأثير هذه التحديات على دقة وجودة الترجمة، ويقترح تطوير مورد لغوي خاص باللغة العربية واستخدام تقنيات الذكاء الاصطناعي المتقدمة لتحسين الترجمة الآلية العربية [24]. توضح هذه الورقة البحثية تطور الذكاء الاصطناعي والتحديات التي تواجهه عند تطبيقه على اللغة العربية كنموذج. ويناقش التعقيدات الخاصة باللغة العربية، بما في ذلك التنوع اللهجي والتراكيب المعقدة للكلمات ومحدودية الموارد اللغوية المطلوبة لتدريب النماذج. كما يوضح الفجوة بين إمكانات الذكاء الاصطناعي وتطبيقه في اللغة العربية، ويقترح حلولاً، مثل بناء قواعد بيانات

لغوية وتشجيع التعاون بين الأوساط الأكاديمية والصناعية لتطوير تقنيات يمكن أن تخدم الناطقين باللغة العربية بفعالية [22] .

تستعرض التطورات الأخيرة في مجال الذكاء الاصطناعي التخاطبي باللغة العربية، وتناقش ايضا التطورات في تقنيات معالجة اللغة الطبيعية وتحليل البيانات الضخمة لتطوير نماذج يمكنها فهم اللهجات العربية والاستجابة لها. كما تتناول التحديات مثل تعقيد اللغة العربية وتنوع اللهجات وصعوبة الحصول على البيانات المناسبة لغرض التدريب. ويعرض الإنجازات في بناء أنظمة حوار ذكية ويقترح تحسينات لتطوير تجارب المستخدمين في التطبيقات العربية [23].

المبحث الثاني

دورخوارزميات الذكاء الاصطناعي في توليد المحتوى العربي

المطلب الاول :مز ايا وعيوب استخدام الخوارزميات الذكية

1.مزايا استخدام الخوارزميات الذكية:

لقد أدت خوارزميات الذكاء الاصطناعي إلى تحسين كبير في عملية توليد المحتوى، خاصة في مجال تحليل النصوص العربية. ومن أبرز فوائدها زيادة الكفاءة في إنشاء المحتوى [2]، حيث يلاحظ الطلاب تأثيرًا إيجابيًا لاستخدام تكنولوجيا الذكاء الاصطناعي على مهاراتهم الكتابية. إذ توفر لهم أدوات متقدمة في القواعد والإملاء، بالإضافة إلى أنظمة للكشف عن الانتحال وتلخيص المحتوى، مما يسهم في تعزيز قدراتهم.

كما أن خوارزميات الذكاء قادرة على معالجة كميات كبيرة من البيانات وبسرعة، مما يجعلها مفيدة في مهام مثل التعرف على الصور والكلام، وهي ميزة حاسمة في تحليل النصوص العربية[1]. وهذا يساعد المؤسسات على تبسيط عمليات إنشاء المحتوى وزيادة الإنتاجية، حيث تتمكن من معالجة البيانات بشكل أسرع من الانسان، بالإضافة إلى استخراج المعلومات القيمة من مجموعات البيانات الكبيرة.

جدول 1: تكرارومدى استخدام الطلاب للذكاء الاصطناعي

الخاصية	مستوى القبول				
	1	2	3	4	5
استخدام تقنيات الذكاء الاصطناعي في كتابة مقالاتي قد عزز قدرتي العامة على الكتابة.	5,3	9,2	20,2	28,1	37,2
التحقق من القواعد والإملاء المدعوم بالذكاء الاصطناعي قد ساعدني في تحديد وتصحيح أخطاء الكتابة، مما ساهم في نمو قدرتي على الكتابة	6,6	12,1	21,3	26,4	33,6
أنظمة الكشف عن الانتحال المعتمدة على الذكاء قد زادت من فهمي لأهمية النزاهة الأكاديمية وقيمة الأصالة في الكتابة.	7,3	8,4	12,1	35,6	36,6
تلخيص المحتوى المولد بالذكاء الاصطناعي قد حسّن قدرتي على استخلاص الأفكار الأساسية من المقالات البحثية الصعبة، مما حسن فهمي للكتابة.	4,8	11,4	16,1	37,4	30,3
استخدام مساعدات الكتابة بالذكاء الاصطناعي قد زاد من وضوح وترابط أعمالي، مما حسن بشكل إيجابي أسلوبي في الكتابة .	5,3	9,4	10,2	35,5	39,6
المخططات التفصيلية المولدة بالذكاء الاصطناعي قد ساعدتني في تنظيم أفكاري	3,5	8,7	13,5	43,2	31,1

بشكل أكثر كفاءة وتحسين تنظيم عملي.					
استخدام أدوات الذكاء الاصطناعي في كتابة مقالاتي قد أجبرني على تقييم ونقد عملي الخاص،مما أدى إلى زيادة قدراتي على التحرير الذاتي	5,7	10,5	15,9	35,8	32,1
تقنيات الترجمة اللغوية المعتمدة على الذكاء الاصطناعي قد قدمت لي مجموعة واسعة من المعلومات الأكاديمية، مما وسع من وجهة نظري ومجموعة مهاراتي الكتابية.	8,5	14,2	17,3	38,7	21,3
تقنيات الذكاء التي توفر اختيارات للكلمات والعبارات المناسبة للسياق قد ساعدتني على توسيع مفرداتي وتحسين جودة كتابتي.	6,8	15,7	14,2	29,6	33,7
أشعر أن استخدام تقنيات الذكاء قد حسّن قدراتي الكتابية وأثر بشكل كبير على كتابة مقالاتي الأكاديمية.	5,3	6,5	8,3	44,5	35,4
مستوى الاتفاق: 1= أعارض بشدة، 2= أعارض، 3= محايد، 4= أوافق، 5= أوافق بشدة					

المصدر: A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis and Marzuki. "Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective". Dec 2023.

جدول 2: تأثير الذكاء الاصطناعي في كتابة المقالات الأكاديمية

الخاصية	مستوى القبول				
	1	2	3	4	5
أقوم بفحص وتعزيز أعمالي باستخدام أدوات تصحيح القواعد والإملاء المدعومة بالذكاء الاصطناعي.	4,6	11,2	21,3	37,5	25,4
أستخدم تقنيات كشف الانتحال المستندة على الذكاء لضمان أصالة كتاباتي الأكاديمية.	2,2	2,4	17,6	32,2	45,6
تلعب الملخصات المولدة بالذكاء الاصطناعي دورًا في مساعدتي على فهم المقالات البحثية الصعبة لتضمينها في كتاباتي.	4,5	16,4	36,7	23,6	18,8
تساعدني تقنيات الذكاء الاصطناعي في توليد الاستشهادات والمراجع بأنواع مختلفة من الصيغ تلقائيًا.	4,5	14,6	24,2	28,2	28,5
تقدم أدوات الكتابة المدعومة بالذكاء الاصطناعي اقتراحات للكلمات والعبارات المناسبة للسياق لمساعدتي في تحسين كتابة مقالاتي.	8,5	18,3	18,3	32,6	22,3
يقترح نظام مدعوم بالذكاء الاصطناعي مواضيع البحث والمصادر المناسبة لمساعدتي في إعداد مقالاتي.	4,2	4,9	11,3	34,5	45,1

تساعدني أدوات الذكاء الاصطناعي في تعديل أسلوب ونبرة مقالاتي لتتناسب مع معايير أكاديمية معينة.	7,5	11,1	18,3	27,3	35,8
أستخدم المخططات التفصيلية المولدة بالذكاء لتنظيم أفكاري بفعالية قبل الكتابة.	5,9	10,8	21,6	33,3	28,4
تساعدني أدوات إدارة الوقت المدعومة بالذكاء في تتبع عملي في كتابة المقالات والالتزام بالمواعيد النهائية.	3,1	4,8	12,2	32,6	47,3

المصدر. A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis and Marzuki. "Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective". Dec 2023.

جدول 3: آراء الطلاب حول تعاون الذكاء الاصطناعي مع الكتاب في المستقبل

الخاصية	مستوى القبول				
	1	2	3	4	5
أرى أنه يجب استخدام الذكاء الاصطناعي كمساعد في الكتابة لمساعدة المؤلفين البشر خلال عملية كتابة المقالات.	3,8	6,3	10,7	34,7	44,5
يجب استخدام الذكاء الاصطناعي لاكتشاف المشكلات المحتملة في القواعد والإملاء، مع اتخاذ المؤلفين البشر القرار النهائي بشأن التصحيحات.	7,3	8,3	12,7	40,3	31,4
أفضّل أن يوفر الذكاء الاصطناعي توصيات وأفكار المحتوى، بينما يحتفظ المؤلفون البشر بالتحكم الإبداعي في اتجاه المقال واستدلاله.	8,1	11,3	12,9	31,6	36,1
يجب أن يساعد الذكاء الاصطناعي المؤلفين البشر من خلال اقتراح مواضيع البحث والمصادر المناسبة، ولكن يجب على الكتاب البشريين القيام بالتحليل والنقد الضروريين.	4,4	9,3	13,4	42,3	30,6
أرى أن المخططات التفصيلية المولدة بالذكاء الاصطناعي يمكن أن تكون مفيدة، ولكن يجب أن يكون للمؤلفين البشر حرية تعديلها وتطويرها لتناسب أسلوبهم في الكتابة.	5,6	7,3	13,4	34,6	39,1
يجب استخدام الذكاء الاصطناعي في تلخيص المحتوى وتقديم المساعدة للمؤلفين البشر في تقليص المواد المعقدة إلى فقرات مختصرة وسهلة الفهم.	7,4	6,1	14,6	43,4	28,5
أشمل الذكاء الاصطناعي للمساعدة في الاستشهادات والتوثيق، ولكن يجب على المؤلفين البشر ضمان صحة وملاءمة المصادر المذكورة.	6,8	8,3	13,1	38,6	33,2
بينما يمكن أن يساعد الذكاء الاصطناعي في الترجمة اللغوية، يجب على المؤلفين البشر التحقق من المحتوى المترجم لضمان السياق والدقة.	7,3	8,5	18,1	39,6	26,5
يجب استخدام الذكاء الاصطناعي لاكتشاف المجالات المحتملة للتحسين خلال عملية التحرير، بينما يقوم المؤلفون البشر بإجراء التعديلات والتحسينات النهائية.	8,1	5,4	13,2	40,2	33,1

اعتقد أن الذكاء الاصطناعي والمؤلفين البشر يجب أن يعملوا معًا لتحقيق أقصى استفادة من كفاءة الذكاء الاصطناعي مع الحفاظ على الأصالة والتفكير النقدي البشري في كتابة المقالات.	5,6	7,3	16,1	35,6	35,4

المصدر: A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis and Marzuki. "Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective". Dec 2023.

2.عيوب استخدام الخوارزميات الذكية

1.2 العجز في الإبداع البشر

أن استخدام الخوارزميات الذكية في تحليل النصوص العربية يؤدي إلى تقليص الإبداع في المحتوى المُنتَج. ورغم فوائدها في اقتراح مواضيع البحث وتقديم توصيات للمحتوى وتوليد المخططات، إلا أن هناك مخاوف بشأن الحفاظ على الإبداع البشري أثناء عملية الكتابة[2].

2.2 تحديات الفروق اللغوية والسياق الثقافي

تواجه الخوارزميات الذكية صعوبات في الحفاظ على الفروق اللغوية والسياق الثقافي في النصوص العربية وتعقيدات اللغة العربية، مثل التشكيل والاختلافات في اللهجات حيث انها تشكل تحديات تؤثر على دقة التعرف على الأحرف وتحليل الكلمات)[8]، [12]). فالطبيعة التعبيرية للعربية تجعل من الصعب على النماذج الذكية فهم العلاقات بين الكلمات [11].

3.2 نقص الدعم والتطوير المستمر

يساهم نقص الالتزام بتعزيز النماذج المخصصة لخصائص اللغة العربية في تقليص فرص الإبداع والابتكار. كما أن غياب الدعم الكافي أو التحديثات غير المنتظمة قد يؤدي إلى ظهور أخطاء غير مفحوصة وانتشار البيانات المضللة، مما يؤثر سلبياً على حرية التعبير في وسائل التواصل الاجتماعية [11].

من هنا، نستنتج أن العجز في الإبداع والعقبات اللغوية والثقافية يبرز أهمية تحقيق توازن بين كفاءة الذكاء الاصطناعي والإشراف لضمان الجودة والمصداقية في تحليل النصوص العربية. لذا، يجب تحسين وتطوير مجموعات تدريبية ونماذج مخصصة للغة العربية للتصدي لهذه التحديات.

جدول 4: مزايا وعيوب التعلم الآلي

عيوب التعلم الآلي	مزايا التعلم الآلي
التحيز :يمكن أن تكون خوارزميات التعلم الآلي منحازة إذا كانت بيانات التدريب غير ممثلة للبيانات الواقعية. يمكن أن يؤدي هذا إلى توقعات وقرارات غير دقيقة.	**الأتمتة** :يمكن لخوارزميات التعليم الآلي أتمتة مختلف المهام مثل تحليل البيانات واتخاذ القرارات والتعرف على الأنماط. يمكن أن يوفر هذا الوقت ويقلل من الحاجة إلى التدخل البشري.
عدم الشفافية :يمكن أن تكون خوارزميات التعلم الآلي صعبة الفهم والتفسير. يمكن أن يجعل هذا من الصعب شرح الأسباب وراء التوقعات والقرارات.	**قابلية التوسع** :يمكن توسيع خوارزميات التعلم الآلي لمعالجة كميات كبيرة من البيانات. يمكن أن يكون هذا مفيدًا بشكل خاص للمهام التي تتطلب الكثير من البيانات، مثل تمييز الصور والكلام.
3. **العمومية المحدودة** :تم تصميم خوارزميات التعليم الآلي لتعمل بشكل جيد	**تحسين الدقة** :يمكن لخوارزميات التعلم الآلي التعلم من البيانات لعمل

على بيانات التدريب ولكن قد لا تعمم بشكل جيد على البيانات الجديدة. يمكن أن يؤدي هذا إلى فرط التكييف وأداء ضعيف على البيانات الواقعية.	توقعات بدقة عالية. يمكن أن يكون هذا مفيدًا بشكل خاص للمهام مثل كشف الاحتيال وتشخيص الأمراض والتنبؤ بالطقس.
الحاجة إلى كميات كبيرة من البيانات :تتطلب خوارزميات التعلم الآلي كميات كبيرة من البيانات للتعلم منها. يمكن أن يكون هذا تحديًا في حاله كون البيانات نادرة أو مكلفة.	4. **المرونة** :يمكن تطبيق خوارزميات التعلم الآلي على أنواع من البيانات مثل النصوص والصوت والصور. يجعلها هذا متعددة الاستخدامات للغاية لتطبيقات مختلفة.
التعقيد الحسابي :يمكن أن تكون بعض من خوارزميات التعلم الآلي معقدة حسابيًا وتتطلب موارد حوسبة كبيرة للتشغيل.	5. **التخصيص** :تستخدام خوارزميات التعلم الآلي لتخصيص التوصيات والخدمات بناءً على تفضيلات وسلوكيات الأفراد.
نقص الفهم السياقي :قد لا تتمتع خوارزميات التعلم الآلي بفهم سياقي للبيانات التي تعمل معها، مما يمكن أن يؤدي إلى توقعات وقرارات غير دقيقة.	6. **اتخاذ القرارات في الوقت الحقيقي** :يمكن استخدام خوارزميات التعلم الآلي لاتخاذ قرارات في الوقت الفعلي بناءً على البيانات، مما يمكن أن يكون مفيدًا بشكل خاص لتطبيقات مثل المركبات الذاتية القيادة وكشف الاحتيال.
المخاوف الأخلاقية :تستخدام خوارزميات التعليم الآلي لأغراض غير أخلاقية، مثل التمييز أو انتهاك الخصوصية.	7.**تقليل الخطأ البشري** :يمكن لخوارزميات التعلم الآلي تقليل خطر الخطأ البشري في اتخاذ القرارات والتحليل.
الاعتماد على (بيانات التدريب) :تعتمد خوارزميات التعلم الآلي على جودة وكمية بيانات التدريب. يمكن أن تؤدي البيانات ذات الجودة المنخفضة أو غير الكافية إلى أداء ضعيف.	8. **التحسين المستمر** :تتمكن الخوارزميات التعلم من البيانات الجديدة لتحسين أدائها ودقتها باستمرار.
استغراق الوقت :يمكن أن يكون تطوير وتدريب خوارزميات التعلم الآلي مستهلكًا للوقت، خاصة بالنسبة للمهام المعقدة.	9.**فعالية التكلفة** :يمكن أن تكون خوارزميات التعلم الآلي فعالة من حيث التكلفة مقارنة بالتحليل اليدوي واتخاذ القرارات.
التفسير المحدود :قد لا توفر خوارزميات التعلم الآلي دائمًا تفسيرات واضحة لتوقعاتها وقراراتها. يمكن أن يجعل هذا من الصعب فهم النتائج والثقة بها.	10. **الميزة التنافسية** :توفر الميزة التنافسية للمنظمات من خلال تمكينها من اتخاذ قرارات أكثر استنارة ومستندة إلى البيانات.

المصدر: A. Omar and T. A. El-Hafeez. "Quantum computing and machine learning for Arabic language sentiment classification in social media". Oct 2023

المطلب الثاني :تأثير استخدام الذكاء الاصطناعي على جودة المحتوى العربي

1.التأثير على جودة ومصداقية المحتوى العربي

1.1تحسين الجودة من خلال التدقيق اللغوي والتحرير الآلي[2]

أدى دمج الخوارزميات الذكية في أدوات التدقيق اللغوي والتحرير الآلي إلى تعزيز مهارات الكتابة والقواعد اللغوية، فضلاً عن تحسين دقة الإملاء لدى الطلاب.

- أهمية الأدوات المعززة بتقنيات الذكاء الاصطناعي

يساهم استخدام الأدوات المدعومة بالذكاء، مثل مدققات القواعد وكاشفات الانتحال وملخصات المحتوى وميزات تنظيم المقالات، بشكل كبير في تحسين أسلوب الكتابة وهيكلها. هذه الأدوات لا تقتصر على تصحيح الأخطاء فحسب، بل تشجع أيضًا على التقييم النقدي للمحتوى، مما يعزز جودة الكتابة الأكاديمية ويضمن أن تكون الأعمال المقدمة أصلية ومنظمة بشكل فعّال.

- تعزيز النزاهة الأكاديمية

ساهم استخدام الذكاء في توليد الاستشهادات وإدارة المراجع وخدمات الترجمة في توسيع الوصول إلى الموارد الأكاديمية، مما أدى إلى تعزيز المفردات وزيادة تعقيد النصوص. نستنتج من ذلك أن استغلال تقنيات الذكاء في تحليل النصوص العربية يُحسن من معايير الجودة والمصداقية، ويعزز مهارات الكتابة، كما يضمن الدقة في الأعمال المكتوبة.

جدول 5: تقييم الجودة للأوراق المشمولة في المراجعة

ت	العنوان	س6	س5	س4	س3	س2	س1	الدرجة	الجودة
1	نهج هجين للتعلم الآلي لتقييم مصداقية صفحات الويب الطبية باللغة العربية	1	0.5	1	1	0.5	1	5	عالية
2	كشف المعلومات الصحية الخاطئة في محتوى الويب: نهج قائم على Web2Vec للكشف عن المعلومات الهيكلية، المعتمدة على المحتوى، والواعية للسياق	1	0.5	1	1	0.5	1	5	عالية
3	موثوق أم لا؟ تصنيف آلي لصفحات الويب حول تطعيم الأطفال الصغار باستخدام التعلم الآلي الخاضع للإشراف	1	0.5	1	0	0	1	3.5	متوسطة
4	التنبؤ بجودة وثائق الويب الصحية باستخدام خصائصها	1	0.5	0.5	1	0	1	4	عالية
5	AutoDiscern تقييم جوده المعلومات الصحية عبر الإنترنت باستخدام شبكات عصبية قائمة على الانتباه التراتبي	1	0.5	1	1	0.5	1	5	متوسطة
6	كشف المعلومات الصحية الخاطئة على الويب الاجتماعي: نظرة عامة ونهج علم البيانات	1	0.5	1	0	0	1	3.5	متوسطة
7	استخدام التعلم الآلي للتعرف التلقائي على المعلومات الصحية المبنية على الأدلة على الويب	1	0.5	1	1	0.5	1	5	عالية
8	تقييم ذكي للمعلومات الصحية عبر الإنترنت لكبار السن	1	0.5	1	0	0.5	1	4	متوسطة
9	Vec4Cred نموذج لكشف المعلومات الصحية الخاطئة في صفحات الويب	1	0.5	1	1	0.5	1	5	عالية

المصدر: Y. K. A. Baqraf, P. Keikhosrokiani and M. Al-Rawashdeh. "Evaluating online health information quality using machine learning and deep learning: A systematic literature review". Nov 2023

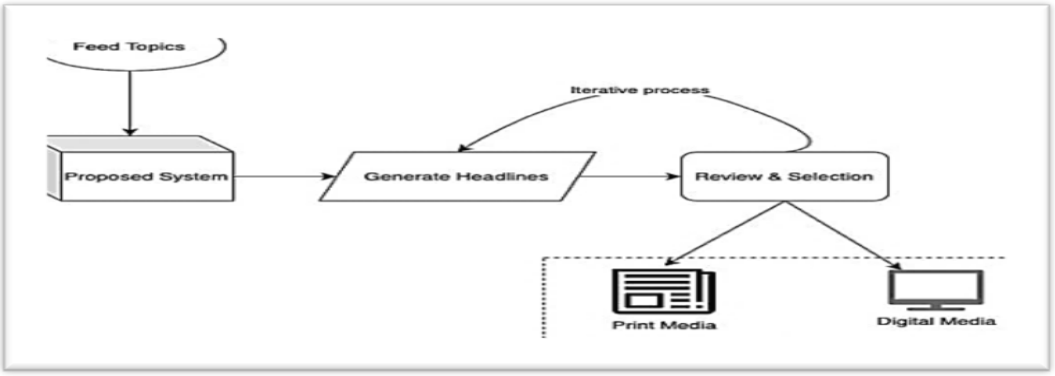
التحليل:

1. الأوراق ذات الجودة العالية: الأوراق 1، 2، 4، 7، و9 حصلت على تصنيف "عالية"، حيث تراوحت درجاتها بين 4 و5.
2. الأوراق ذات الجودة المتوسطة: الأوراق 3، 5، 6، و8 حصلت على تصنيف "متوسطة"، ودرجاتها تتراوح بين 3.5 و5.
3. أعلى الدرجات: الأوراق 1، 2، 5، 7، و9 حققت أعلى درجة، وهي 5.
4. أدنى الدرجات: الأوراق 3 و6 حصلت على أدنى درجة، وهي 3.5.
5. العوامل المشتركة: جميع الأوراق ذات الدرجات العالية تتمتع بتقييمات مرتفعة باستمرار في س1، س4، وس6. بينما الأوراق ذات الجودة المتوسطة تظهر تباينًا أكبر، خاصة في س2 وس3.

الملاحظات الرئيسية:

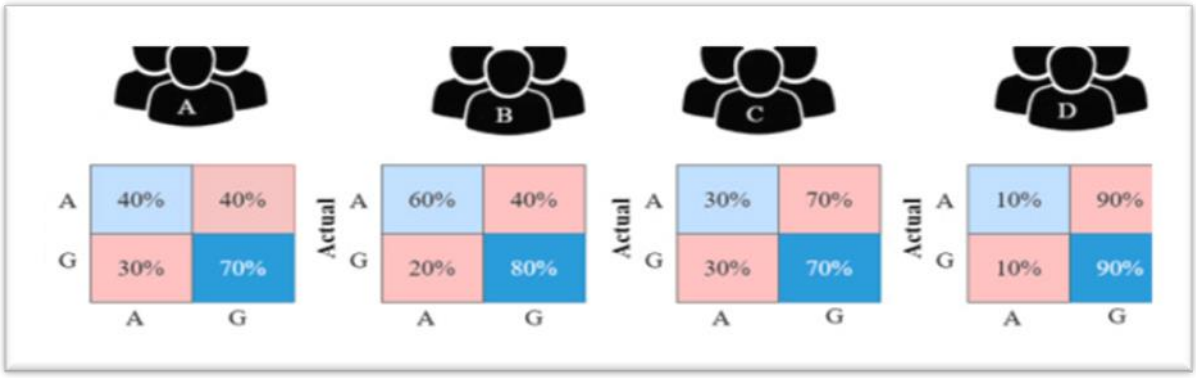
• الأوراق المصنفة كـ "عالية" تميل إلى الحصول على درجات مرتفعة باستمرار في معظم المعايير، وخاصة في المعايير س1، س4، وس6. في المقابل، الأوراق ذات الجودة "متوسطة" تظهر تباينًا أكبر، خاصة في المعايير س2 وس3، مما قد يشير إلى جوانب تكون فيها هذه الأوراق أقل قوة.

الشكل 1: استخدام التطبيق للنظام المقترح



المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

الشكل 2: أداء كل مجموعة في اللغات ذات الموارد المحدودة بعد تنفيذ HPPE



المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

جدول 6: نتائج القرارات البشرية من لغة منخفضة المستوى تم الحصول عليها من تطبيقاتHPPE.

التصنيف	المحتوى المنشاء (المتولده)	المحتوى الاصليه
المحتوى الاصلية المتوقع	58%	42%
المحتوى المتولد المتوقع	72%	28%

المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

جدول 7: الأنماط الرئيسية الخمسة لتصنيف الكلمات المستخرجة من العناوين المولدة الإنجليزية التي نجحت في اختبار تورينج

ت	نمط تصنيف الكلمات	مثال على العنوان
1	NNP VBD DT NN	Apple launched the iPhone
2	DT JJ NN VBZ JJ	The new policy is effective
3	NN IN DT NN	Growth in the sector
4	DT NN VBD IN NN	The company expanded in 2023
5	NNP VBZ RB JJ	Tesla is highly innovative

المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

المفتاح:

- NNP: اسم علم، مفرد

RB: ظرف

- VBD: فعل بصيغة الماضي

JJ: صفة

VBZ: فعل بصيغة المضارع للشخص الثالث المفرد

- DT: أداة تعريف أو تنكير

IN: حرف جر أو أداة ربط

NN: اسم، مفرد أو غير معدود

جدول 8: مثال تطبيقي يوضح الأنماط الرئيسية الخمسة لتصنيف الكلمات المستخرجة من العناوين المولدة الإنجليزية التي نجحت في اختبار تورينج.

antecedent	Consequent	Support	Confidence	Lift
{VERB, NOUN}	{PROPN}	0.3	0.33	1.11
{VERB}	{NOUN, PROPN}	0.3	0.33	1.11
{ADP, PROPN}	{VERB}	0.13	1	1.11
{ADP, NOUN, PROPN}	{VERB}	0.13	1	1.11
{VERB, NOUN}	{ADP, PROPN}	0.13	0.15	1.11

المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

النمط السابق: هو النمط التصنيفي للكلمات الذي يأتي قبل النمط اللاحق.

النمط اللاحق: هو النمط التصنيفي للكلمات الذي يأتي بعد النمط السابق.

الدعم: هو نسبة ظهور هذا التسلسل في العناوين المولدة.

الثقة: هي نسبة النمط اللاحق بالنسبة للنمط السابق.

جدول 9: الأنماط الرئيسية الخمسة لتصنيف الكلمات المستخرجة من العناوين المولدة الإنجليزية التي لم تنجح في اجتياز اختبار تورينج

ntecedent	Consequent	Support	Confidence	Lift
{PROPN, VERB}	{ADP}	0.29	1	1.75
{ADP, VERB}	{PROPN}	0.29	0.5	1.75
{PROPN, NOUN}	{ADP}	0.29	1	1.75
{PROPN, NOUN, VERB}	{ADP}	0.28	1	1.75
{ADP, NOUN, VERB}	{PROPN}	0.28	0.5	1.75

المصدر: N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023.

جدول 10: مقايس التقييم

مقياس التقييم	الوصف	معادلة
الدقة (Accuracy)	هو إحصاء يلخص مدى أداء النموذج في جميع الفئات. يكون مفيدًا عندما تكون جميع أنواع الفئات ذات أهمية متساوية. يتم حسابه كنسبة بين عدد الأحكام الصحيحة إلى العدد الإجمالي للأحكام.	(TP+TN)/(TP+TN+FP+FN)
الدقة (Precision	تقيس مدى دقة النموذج في تصنيف العينة على أنها إيجابية. يتم تحديدها كنسبة عدد العينات الإيجابية المصنفة بشكل صحيح إلى العدد الإجمالي للعينات الإيجابية (سواء كانت مصنفة بشكل صحيح أو غير صحيح	TP/(TP+FP)
الاسترجاع (Recall)	قدرة النموذج على تحديد عينات ايجابية. يتم حسابه عن طريق تقسيم عدد العينات الإيجابية التي تم تصنيفها بشكل صحيح كإيجابية على العدد الإجمالي للعينات الإيجابية.	TP/(TP+FN)
مقياس F (F-measure	يتم تحديده عن طريق حساب المتوسط التوافقي للدقة	(2*Precision*Recall)/(Precision+

	والاسترجاع.	Recall) = (2*TP)/((2*TP)+FP+FN)
حساسية (Sensitivity)	النسبة المئوية للإيجابيات الحقيقية المكتشفة بشكل صحيح وتقدر مدى فعالية التنبؤ بالفئة الإيجابية.	TP/((TP+FN))
النوعية (Specificity)	هي المكمل للحساسية، وهي معدل الصحيح السلبي الذي يلخص مدى فعالية التنبؤ بالفئة السلبية. قد تكون حساسية التصنيف المتحيز أكثر أهمية من النوعية.	TN/(FP+TN)
المتوسط الهندسي Geometric-mean (G-mean))	هو مقياس يجمع بين الحساسية والنوعية في قيمة واحدة توازن بين الهدفين.	

المصدر: P. Nandwani and R. Verma. "A review on sentiment analysis and emotion detection from text". Jul 2021.

TP (الإيجابيات الحقيقية): تشير بانها (إيجابية بشكل صحيح).

TN (السلبيات الحقيقية): تشير بانها (سلبية بشكل صحيح). .

FP (الإيجابيات الكاذبة): تشير بشكل (خاطئ على أنها إيجابية).

FN (السلبيات الكاذبة): تشير بشكل (خاطئ على أنها سلبية).

المطلب الثالث: المخاطر المحتملة الناتجة عن نشر المعلومات المضللة أو المتحيزة

1.خطوره نشر المعلومات المتحيزة أو المضللة

تُستخدم خوارزميات الذكاء في تحليل النصوص العربية، مما يثير مخاوف بشأن انتشار المعلومات المتحيزة اوالمضللة [8]. يعتبر التواصل الاجتماعي مصدراً رئيسياً للأخبار، مما يؤدي إلى إنتاج كميات هائلة من البيانات. ومع ذلك، فإن سهولة النشر تُسهم في تداول الأخبار الكاذبة والمضللة، خاصةً في ظل الخصائص الفريدة للغة العربية، مثل الغموض والتنوع اللهجي، التي تعقد تحليل النصوص على هذه المنصات.

علاوة على ذلك، [11] لم يُعطِ مطورو الذكاء الاصطناعي الأولوية الكافية لتحسين النماذج لفهم النصوص العربية، مما قد يؤدي إلى أخطاء ورقابة غير مقصودة تؤثر على حرية التعبير، أو تساهم في نشر المعلومات المضللة أو خطاب الكراهية. كما أن الفروقات في التعديل الخوارزمي بين المناطق واللهجات المختلفة [11]يمكن أن تؤدي إلى سوء الفهم وإجراءات غير مناسبة. على سبيل المثال، قد تتجنب الصفحات المرتبطة بالمنظمات المتطرفة التعديل من خلال استخدام الرموز التعبيرية.

كذلك، وُجهت انتقادات للتعديل الآلي للمحتوى على وسائل التواصل الاجتماعي، حيث تم رصد حالات إغلاق حسابات نشطاء أو صحفيين دون أسباب واضحة، مما يثير القلق بشأن دقة وحيادية أدوات الذكاء المستخدمة في التعديل[11].

يُظهر دمج خوارزميات الذكاء في تحليل النصوص العربيه وجود مخاطر تتعلق بنشر المعلومات المضللة والمتحيزة، مما يتطلب إشرافًا دقيقًا عند استخدام هذه التقنيات.

2.قواعد البيانات والسجلات المستخرجة.

جدول 11: أسماء قواعد البيانات، عبارات البحث، والسجلات المستخرجة

ت	اسم قاعدة البيانات	السجلات المستخرجة	عبارة البحث	الحقول والفلاتر
1	الرقمية ACM مكتبة	180	B	2017–2022
2	ساينس دايركت	346	C	مقالات بحثية، ملخصات المؤتمرات، 2017–2022
3	سكوبس	2212	A	الإنجليزية، مقالات بحثية، أوراق المؤتمرات، 2017–2022
4	ويب أوف ساينس	42	A	2017–2022
5	سبر ينجر لينك	228	A	الإنجليزيه، التخصص: الطب والصحة العامة، التخصص الفرعي: المعلوماتية الصحية، مقالات، 2017–2022
6	مكتبة وايلي الرقمية	41	B	مجلات، علوم الكمبيوتر، 2017–2022
7	بوب ميد	128	A	2017–2022
8	إميرالد إنسايت	101	B	2017–2022
9	تتبع الاستشهادات	1		2017–2022

المصدر: Y. K. A. Baqraf, P. Keikhosrokiani and M. Al-Rawashdeh. "Evaluating online health information quality using machine learning and deep learning: A systematic literature review". Nov 2023

وصف الخصائص

جدول 12: وصف كل ميزة

الوصف	النوع	الخاصية
يحتوي هذا العمود على تصنيف جنس المستخدم. يحتوي على القيمتين "F" للإناث و "M"للذكور. يتم الحصول على القيمة بملاحظة اسم الحساب، الصورة المعروضة، أو من البيانات النصية في محتوى المنشورات.	اسمي	الجنس
تحتوي هذه الخاصية على قيم مكونة من سلاسل تم تصفيتها وتحويلها إلى كلمات متجهة. محتوى العمود مفيد لتصنيف الآراء حيث يتم استنتاج المصطلحات (الإيجابية والسلبية) للوصول إلى قرار التصنيف سواء كان سلبياً أو إيجابياً.	اسمي	النص
يتعلق هذا العمود بنتائج تصنيف الحالة. يتم تمثيل الفئة بقيمتين اسميتين: تصنيف إيجابي وتصنيف سلبي. يتم استنتاج نتيجة التصنيف من خلال تحليل الكلمات والعبارات المستخدمة في عمود النص.	اسمي	الفئة

المصدر: B. A. sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi and Sunday O. Olatunji. "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms". Feb 2022.

3. تحقيق التوازن بين الكفاءة والدقة

1.3. تحقيق التوازن بين السرعة والدقة في توليد النصوص

ضمان دقة النصوص العربية التي تنتجها خوارزميات الذكاء الاصطناعي يتطلب اتخاذ تدابير فعالة لمراقبة الجودة. كما أظهرت الأبحاث المتعلقة بضبط المعلمات الفائقة لخوارزميات التعلم الآلي ([6])، فإن اختيار هذه المعلمات يؤثر بشكل كبير على أداء ودقة الخوارزميات. يتم ذلك من خلال تقنيات مثل البحث الشبكي، البحث العشوائي، تحسين بايزي، تحسين سرب الجسيمات (PSO)، والخوارزميات الجينية (GA). وقد تم تحديد المعلمات الفائقة المثلى لمصنفات متنوعة مثل SVM، الانحدار اللوجستي، نايف بايز، شجرة القرار، ونماذج الغابات العشوائية.

فيما يتعلق بتصنيف الأخبار العربية بناءً على ضبط المعلمات ([5])، أظهرت المقارنات أن البحث العشوائي يتفوق على البحث الشبكي من حيث الدقة، Precision، Recall، ودرجة F1 في خوارزميات مثل الانحدار اللوجستي متعدد الحدود (MLR) وSVM والشبكات العصبية الاصطناعية (ANN). وهذا يبرز أهمية اختيار المعلمات الفائقة المناسبة لتحسين دقة النماذج.

علاوة على ذلك، تعتبر تقنيات المعالجة المسبقة ضرورية لتحليل بيانات النصوص العربية على وسائل التواصل الاجتماعي ([4])، حيث تسهم معالجة القيم المفقودة وتصفية الضوضاء في البيانات على منصات مثل Instagram وSnapchat وTwitter في تعزيز موثوقية نتائج تحليل المشاعر.

كما توفر الطرق القائمة على القواعد، مثل خوارزمية تحسين التوازن الثنائي ([3])، رؤى حول تحقيق توازن بين الدقة وقابلية التفسير. بالإضافة إلى ذلك، تساعد مقارنة تقنيات التصنيف مثل SVM وK-NN ونايف بايز مع المنهجيات ذات الصندوق الأبيض مثل PART وRIPPER في تقييم التبادلات بين الدقة والشفافية في تحليل النصوص.

جدول 13: عدد مجموعات البيانات قبل وبعد مرحلة المعالجة المسبقة

نوع المنصة	قبل مرحلة المعالجة المسبقة	بعد مرحلة المعالجة المسبقة
Instagram	7538	514
Snapchat	1932	284
Twitter	1452	462

المصدر: B. A. sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi and Sunday O. Olatunji. "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms". Feb 2022.

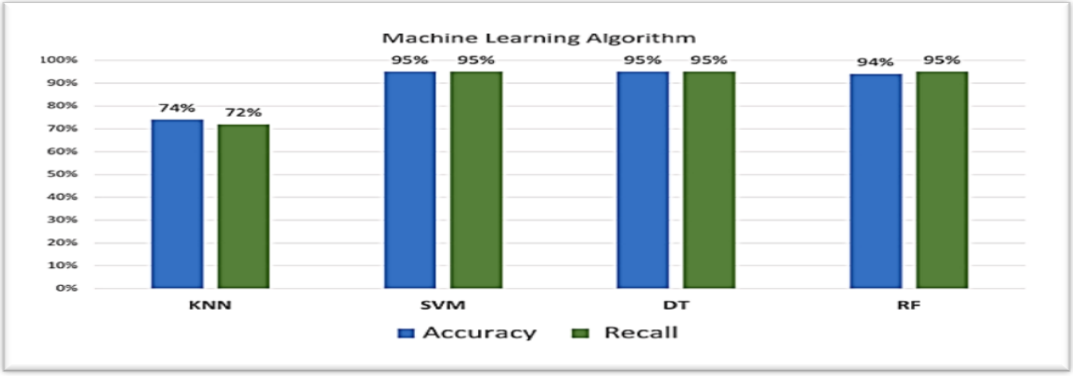
جدول 14: مقارنة النهج المقترح مع بعض طرق التصنيف الأخرى

رقم	نوع النهج	النهج	الدقة(%)	الاسترجاع(%)	الدقة(Precision %)	حجم المصنف
1	النهج الأسود	SVM	69.20	–	–	–
2		K-NN	50.00	–	–	–

3		NB	83.2	83.2	83.2	–
4	النهج الأبيض	PART	73.4	73.5	73.4	10
5		RIPPER	72.8	73	72.8	7
6		OneR	61	73.1	61	–
7		C4.5	68.6	68.7	68.6	41
8		REPTree	69.2	69.4	69.2	29
9		نهجنا	84.0	68	100	13

المصدر: H. Rahab, H. Haouassi and A. Laouid. "Rule-Based Arabic Sentiment Analysis using Binary Equilibrium Optimization Algorithm". Sep 2022.

الشكل 3: مقارنة بين مصنِّفات مختلفة بناءً على الدقة والاسترجاع.



المصدر: O. Alsemaree, Atm S. Alam, S. S. Gill and S. Uhlig. "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions". May 2024

جدول 15: تحليل مقارن للمشاعر لكل مصنف

الخوارزميه	الصنف	الدقه	الدقه	الاسترجاع	F1
KNN	−1	74.0	74.0%	72.0	73.0
	0	74.0		71.0	72.0
	+1	73.0		79.0	76.0
SVM	−1	97.0	95.0%	95.0	95.0
	0	94.0		95.0	96.0
	+1	95.0		95.0	95.0
DT	−1	96.0	95.0%	95.0	95.0
	0	93.0		93.0	93.0
	+1	93.0		95.0	94.0

96.0	95.0	97.0		−1	RF
94.0	95.0	93.0	94.0%	0	
95.0	95.0	96.0		+1	

المصدر: O. Alsemaree, Atm S. Alam, S. S. Gill and S. Uhlig. "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions". May 2024

المبحث الثالث

الفرص والتحديات والاتجاهات المستقبل

المطلب الاول :الفرص في صناعة المحتوى العربي

1.توسيع نطاق السوق من خلال إنشاء محتوى متعدد اللغات:

- دمج الذكاء الاصطناعي: تسهم خوارزميات الذكاء في تحسين الكتابة والفهم، واكتشاف الأخطاء في المقالات الأكاديمية، كما تساعد في تنظيم الأفكار وإثراء المفردات، مما يمثل نقلة نوعية في صناعة المحتوى العربي.
- جمهور أوسع: يتيح إنشاء محتوى متعدد اللغات لمطوري المحتوى العربي الوصول إلى جمهور أكبر وتلبية تفضيلات لغوية متنوعة.
- تحليل المشاعر: تعتبر أدوات تحليل المشاعر ضرورية لفهم كيفية إدراك الجمهور ومراقبة سمعة العلامات التجارية.
- ضبط المعلمات: تقنيات مثل ضبط المعلمات (Hyperparameter Tuning) تعزز من تصنيف الأخبار العربية، مما يزيد من دقة عمليات إنشاء المحتوى متعدد اللغات.
- الأدوات المتخصصة: تطوير أدوات مخصصة لتحليل النصوص العربية، بما في ذلك تقنيات تعلم الآلة المتعمق وأنظمة تصنيف الوثائق، سيدعم جهود إنشاء المحتوى متعدد اللغات.
- الآفاق المستقبلية: تتيح التحسينات المستمرة في تكنولوجيا الذكاء فرصًا كبيرة للنمو والابتكار في قطاع المحتوى العربي.

2. تطوير أدوات وتقنيات جديدة لتحليل النصوص العربية

- النماذج المدربة مسبقاً: تم تعديل نماذج مثل BERT وGPT لتناسب اللغة العربية، مما يعزز من فعالية معالجة النصوص العربية في مختلف مهام معالجة اللغة الطبيعية.
- النماذج متعددة اللغات: تساهم هذه النماذج في تحسين دقة مهام كالترجمة الآلية للنصوص العربية وتحليل المشاعر.
- نماذج اللهجات: تم إنشاء نماذج جديدة تلبي احتياجات اللهجات العربية المتنوعة، وهو امر مهم في السياقات غير الرسمية ووسائل التواصل الاجتماعي.
- التحليلات المتقدمة: تعمل التطورات في تحليل المشاعر، واكتشاف العواطف، والتعرف على الكيانات للنصوص العربية تقدماً ملحوظاً، مما يحسن من التطبيقات مثل مراقبة وسائل التواصل الاجتماعي وتحليل آراء العملاء.
- التعرف على الصوت: تسهم التحسينات في تقنيه تمييز الصوت العربي وأنظمة تحويل النص إلى كلام في تعزيز التفاعل الصوتي باللغة العربية.

المطلب الثاني: تحديات خوارزميات الذكاء الاصطناعي

1.التحديات التي تواجهها الخوارزميات في المحتوى العربي

-تعقيد اللغة: يواجه الذكاء تحديات نتيجة تركيب اللغة العربية واختلافاتها السياقية وتعقيدات الكتابة.

-قضايا اللهجات والمجموعات البيانية: تؤدي اللهجات المتنوعة والمجموعات البيانية المحدودة إلى تعقيد عملية التجزئة والتعرف والتحليل.

-صعوبات معالجة اللغة الطبيعية: تؤثر مشكلات مثل غموض ثنائية اللغة والجمع غير المنتظم على مهام معالجة اللغة الطبيعية واسترجاع المعلومات.

-الاهتمامات الأخلاقية: تثير إدارة المحتوى المدعومة بالذكاء قضايا أخلاقية تتعلق بالتحيز ومشاكل الموثوقية، مما يؤثر على ثقة المستخدمين.

التوصيات لمواجهة التحديات

•حلول مبتكرة: تصميم خوارزميات ذكاء اصطناعي تأخذ بعين الاعتبار الخصائص اللغوية الفريدة للغة العربية.

•التحضير المتقدم: تطبيق تقنيات مثل التجزئة، التصريف، وتحديد الكيانات لتعزيز دقة أداء الذكاء الاصطناعي.

يقدم هذا العرض هيكلًا واضحًا لفهم الفرص والتحديات في صناعة المحتوى العربي، مما يسلط الضوء على كيفية دفع الابتكارات في مجال الذكاء للنمو مع معالجة التعقيدات الخاصة باللغة العربية.

جدول 16: عدد المشاعر الايجابية والسلبية عند إعادة التقييم

المنصة	التوازن	التكرار الزائد	التكرار الناقص
إنستغرام	342إيجابي، 172 سلبي	342إيجابي، 342 سلبي	170إيجابي، 170 سلبي
سناب شات	256إيجابي، 28 سلبي	256إيجابي، 256 سلبي	28إيجابي، 28 سلبي
تويتر	260إيجابي، 202 سلبي	260إيجابي، 260 سلبي	200إيجابي، 202 سلبي

المصدر: B. A. sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi and Sunday O. Olatunji. "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms". Feb 2022.

جدول 17: المعلمات المثلى لنموذج SVM

المنصة	التكرار الناقص	التكرار الزائد	التوازن
إنستغرام	10-fold (نعم)، 70% (نعم)	10-fold (لا)، 70% (لا)	10-fold (لا)، 70% (لا)
سناب شات	10-fold (لا)، 70% (نعم)	10-fold (لا)، 70% (لا)	10-fold (لا)، 70% (لا)
تويتر	10-fold (نعم)، 70% (نعم)	10-fold (نعم)، 70% (لا)	10-fold (لا)، 70% (لا)

المصدر: B. A. sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi and Sunday O. Olatunji. "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms". Feb 2022.

جدول 18: المعلومات الديموغرافية للمشاركين (N = 245)

البند	القيمة	التكرار	النسبة المئوية
الجنس	ذكر	117	47.75%
	أنثى	128	52.25%
فئة العمر	16–18	49	20%
	19–21	147	60%
	22–25	30	12.25%
	أكبر من 26	19	7.75%
نوع الجامعة	حكومية	107	43.67%
	خاصة	138	56.33%
التخصص	اللغة الإندونيسية	107	43.67%
	تعليم اللغة الإنجليزية	73	29.80%
	اللغة الألمانية	43	17.55%
	اللغة العربية	22	8.97%
موقع الجامعات	سولاويسي الجنوبية	45	18.37%
	سولاويسي الوسطى	77	31.43%
	سولاويسي الشمالية	32	13.06%
	غورونتالو	23	9.39%
	مالوكو	12	4.89%
	بالي	39	15.92%
	بابوا	17	6.94%

المصدر: A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis and Marzuki. "Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective". Dec 2023

الشكل 4: تقنيات معالجه اللغات الطبيعية المختلفة المطلوبة للغة العربية

المصدر: N. V. Otten. "Products & Services Document text extraction Online Tool To Extract Text From PDFs & Images". Feb 2024.

2.التداعيات الأخلاقية لاستخدام الذكاء الاصطناعي في توليد النصوص

تثير التداعيات الأخلاقية لاستخدام الذكاء في توليد النصوص باللغة العربية قضايا مهمة. فقد أظهرت الدراسات أن الخوارزميات قد تؤدي إلى تباين في ممارسات التعديل، مما يسبب عدم اليقين بين المستخدمين العرب ويقلل من مشاركتهم في المناقشات السياسية، كما يضعف ثقتهم في أدوات الذكاء الاصطناعي ([11]، [18]). لذا، تبرز أهمية المساءلة والشفافية في التقنيات الخوارزمية لتفادي التمثيل غير المتناسب للعرب والمسلمين في قوائم المنظمات الخطيرة، مما قد يؤدي إلى قيود على التفاعل وإفراط في تعديل المحتوى.

بالإضافة إلى ذلك، فإن التدريب غير الكافي للمشرفين قد يسفر عن تمييز ضد المحتوى غير الإنجليزي، مما يستدعي معالجة تحيزات الشركات التكنولوجية ([11] [14]). ومع تقدم الذكاء الاصطناعي، تزداد المخاوف من التحيز ونشر المعلومات المضللة، مما يتطلب الالتزام بمبادئ الأخلاق مثل العدالة والشفافية والخصوصية لتقليل المخاطر المحتملة ([11]. [18]).

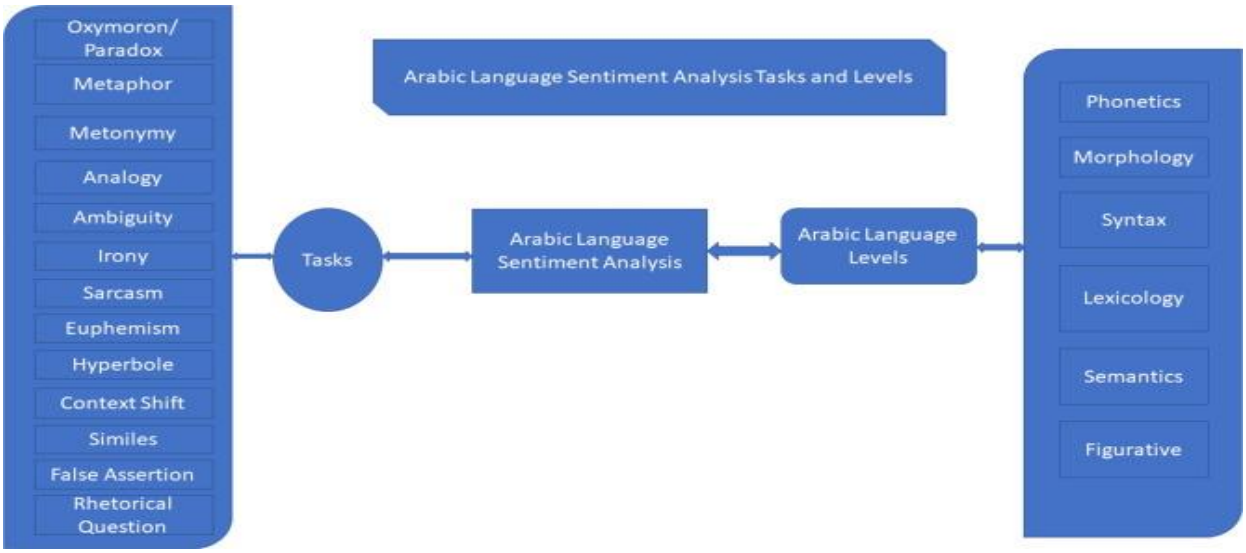
المطلب الثالث: الاتجاهات المستقبلية للذكاء الاصطناعي في المحتوى العربي

1- الاتجاهات والتطورات المستقبلية في تكنولوجيا الذكاء الاصطناعي في المحتوى العربي:

شهدت تكنولوجيا الذكاء الاصطناعي للمحتوى العربي تقدمًا ملحوظًا، على الرغم من التحديات مثل التباينات الإملائية والاختلافات في اللهجات. حيث تسهم النماذج اللغوية المدربة مسبقًا والنماذج متعددة اللغات في تحسين دقة تحليل وتوليد النصوص. كما تعزز تقنيات التعرف على الكلام وتحويل النص إلى كلام من التفاعل الصوتي باللغة العربية ([10]، [8]).

أظهرت الأبحاث في تحليل المشاعر باستخدام مصنفي تعلم الآلة نتائج إيجابية في فهم العواطف والآراء في النصوص العربية. ومن الضروري التركيز على إثراء البيانات المعلّمة وتطبيق تقنيات التعلم بالنقل لتحسين أداء النماذج. هذا سيفتح آفاقًا جديدة في مجالات التحليل النصي الآلي، وتصنيف المشاعر، والترجمة الآلية، مما يعزز التواصل بين الناطقين بالعربية على مستوى عالمي [12].

الشكل 5: مهام ومستويات ALSA



المصدر: A. Alsayat and N. Elmitwally. "A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications)". Mar 2020.

الاستنتاج

تُظهر الدراسات أهمية اعتماد خوارزميات الذكاء الاصطناعي لتحسين دقة وفعالية تحليل النصوص العربية، لاسيما على منصات التواصل الاجتماعية، وتشمل الفوائد استخدام تقنيات كتقسيم النص وإزالة الكلمات الشائعة، بالإضافة إلى نماذج معالجات للغات الطبيعية المتطورة. تعمل تقنيات استخراج الميزات مثل TF-IDF وMRMR، إلى جانب خوارزميات تعلم الآلة مثل KNN وSVM وDT وRF، في تحسين دقة تحليل المشاعر. ومع ذلك، تواجه هذه العملية تحديات مثل عدم توازن توزيع البيانات وانتشار المعلومات المضللة، مما يستدعي ضرورة تبني استراتيجيات لضبط الجودة. من المهم تحقيق توازن بين السرعة والدقة في توليد النصوص مع الحفاظ على سلامة البيانات، وتطوير أدوات مخصصة لتحليل النصوص العربية. تقدم تكنولوجيا الذكاء الاصطناعي إمكانيه كبيرة لتحسين تحليل وإنتاج المحتوى العربي.

المصادروالمراجع:

[1]A. Omar and T. A. El-Hafeez. "Quantum computing and machine learning for Arabic language sentiment classification in social media". Oct 2023. [Online]. Available: https://www.nature.com/articles/s41598-023-44113-7

[2]A. R. Malik, Y. Pratiwi, K. Andajani, I. W. Numertayasa, S. Suharti, A. Darwis and Marzuki. "Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective". Dec 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666374023000717

[3]H. Rahab, H. Haouassi and A. Laouid. "Rule-Based Arabic Sentiment Analysis using Binary Equilibrium Optimization Algorithm". Sep 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9513016/

[4]B. A. sari, R. Alkhaldi, D. Alsaffar, T. Alkhaldi, H. Almaymuni, N. Alnaim, N. Alghamdi and Sunday O. Olatunji. "Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms". Feb 2022. [Online]. Available: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00568-5

[5]I. Jamaleddyn, R. E. ayachi and M. Biniz. "An improved approach to Arabic news classification based on hyperparameter tuning of machine learning algorithms". Jun 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2307187723000627

[6]E. Elgeldawi, A. Sayed, Ahmed R. Galal and Alaa M. Zaki. "Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis". Nov 2021. [Online]. Available: https://www.mdpi.com/2227-9709/8/4/79

[7]N. V. Otten. "Products & Services Document text extraction Online Tool To Extract Text From PDFs & Images". Feb 2024. [Online]. Available: https://spotintelligence.com/2023/10/29/arabic-nlp/

[8]S. AL-Sarayreh, A. Mohamed and K. Shaalan. "Challenges and Solutions for Arabic Natural Language Processing in Social Media". Jul 2023. [Online]. Available: https://www.researchgate.net/publication/372096703_Challenges_and_Solutions_for_Arabic_Natural_Language_Processing_in_Social_Media

[9]A. H. Doohee. "The Contribution of Artificial Intelligence to Learning the Arabic Language". May 2024. [Online .[
Available: https://ej-lang.org/index.php/ejlang/article/view/120

[10]A. Alsayat and N. Elmitwally. "A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications)". Mar 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110866519300945

[11]M. Elswah. "Does AI Understand Arabic? Evaluating the Politics Behind the Algorithmic Arabic Content Moderation". Jan 2024. [Online]. Available: https://www.hks.harvard.edu/sites/default/files/2024-01/24_Elswah_0.pdf

[12]M. SalahEldin, K. M. Mahmoud and H. Kang. "Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey". Dec 2023. [Online]. Available: https://arxiv.org/html/2312.11812v1

[13]S. M. AL-Ghuribi and S. A. M. Noah. "ARABIC TEXT MINING". Nov 2022. [Online]. Available: https://arxiv.org/pdf/2211.02772

[14]B. Katz. "The Intelligence Edge: Opportunities and Challenges from Emerging Technologies for U.S. Intelligence". Apr 2020. [Online]. Available: https://www.csis.org/analysis/intelligence-edge-opportunities-and-challenges-emerging-technologies-us-intelligence

[15]N. Fatima, S. M. Daudpota, Z. Kastrati, A. S. Imran, S. Hassan and N. S. Elmitwally. "Improving news headline text generation quality through frequent POS-Tag patterns analysis". Oct 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197623009028

[16]O. Alsemaree, Atm S. Alam, S. S. Gill and S. Uhlig. "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions". May 2024. [Online .[
Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11070797/

[17]Hamzah, Ali A., W. Alosaimi, H. Saleh, A. Alharb, N. El-Rashidy, A. Elaraby and S. Mostafa. "Frontiers in Artificial Intelligence". Jul 2024. [Online]. Available: https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1408845/full

[18]E. Bird, J. Fox-Skelly, N. Jenner, R. Larbey, E. Weitkamp and A. Winfield. "The ethics of artificial intelligence: Issues and initiatives". Mar 2020. [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

[19]E. Elgeldawi, A. Sayed, Ahmed R. Galal and Alaa M. Zaki. "Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis". Nov 2021. [Online]. Available: https://www.researchgate.net/publication/356292529_Hyperparameter_Tuning_for_Machine_Learning_Algorithms_Used_for_Arabic_Sentiment_Analysis

[20]Y. K. A. Baqraf, P. Keikhosrokiani and M. Al-Rawashdeh. "Evaluating online health information quality using machine learning and deep learning: A systematic literature review". Nov 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10664453/

[21]P. Nandwani and R. Verma. "A review on sentiment analysis and emotion detection from text". Jul 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8402961/

[22] Khalatia, M. M., & Al-Romanyb, T. A. H. (2020). Artificial intelligence development and challenges (Arabic language as a model). Artificial Intelligence, 13(5), 83-105. [Online]. Available: https://www.ijicc.net/images/vol_13/Iss_5/13578_Khalati_2020_E_R.pdf

[23]Fuad, A., & Al-Yahya, M."Recent developments in Arabic conversational AI: a literature review". IEEE Access, 10, 23842-23859, 2022 .[Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9723043

[24]Ali, M. A. "Artificial intelligence and natural language processing: the Arabic corpora in online translation software. International Journal of Advanced and Applied Sciences", 3(9), 59-66,2016 .[Online]. Available: https://www.academia.edu/29353299/Artificial_intelligence_and_natural_language_processing_the_Arabic_corpora_in_online_translation_software

[25]Gardner, H. E."Multiple intelligences: New horizons in theory and practice". Basic books.2008.[Online]. Available : https://www.google.iq/books/edition/Multiple_Intelligences/7Gw4DgAAQBAJ?hl=id&gbpv=1&dq=inauthor:%22Howard+E+Gardner%22&printsec=frontcover