أدوات الكشف عن محتوى الذكاء الاصطناعي ومدى فاعليتها في الكشف عن الانتحال العلمي في البيئة الأكاديمية: دراسة استكشافية

إعـــداد

د/ محمد أحمد ثابت

أستاذ المكتبات المساعد

كلية الآداب - جامعة أسيوط

Email: moh_art85@yahoo.com

DOI: 10.21608/aakj.2024.313266.1847

تاريخ الاستلام: ١٨/ ٨/ ٢٠٢٤م

تاريخ القبول: ١٥/ ٩/ ٢٠٢٤م

ملخص

سعت الدراسة إلى التعريف بأدوات الكشف عن محتوى الذكاء الاصطناعي من حيث ملامحها العامة ونظم تسعيرها وإيجابياتها وسلبياتها، وتقييم مدى فاعليتها ودقتها في الكشف عن الانتحال العلمي في البيئة الأكاديمية، وفي سبيل ذلك اعتمد الباحث على المنهج الوصفي التحليلي، وتطبيقه على عينة بلغ عدد مفرداتها ٨ أدوات، التي بدورها تم استخدامها مع عدة مقالات مختارة "محتوى بشري، محتوى ذكاء اصطناعي" بلغ عددها ٤ مقالات متخصصة في مجال المكتبات والمعلومات، من خلال قياس ملامح تمييز أدوات الكشف AI Detectors بين المقالات المكتوبة بواسطة أدوات الذكاء الاصطناعي والمقالات المكتوبة بواسطة البشر، وكذا قياس دقة تعرف أدوات الكشف AI Detectors على مقالات الذكاء الاصطناعي الخاضعة للترجمة، ومقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة وكان من أهم نتائجها أن أدوات الكشف عن محتوى الذكاء الاصطناعي بمواصفاتها الحالية لا تبدو جاهزة بما يكفي للكشف الدقيق والمقنع عن المحتوى الناتج عن أدوات الذكاء الاصطناعي، بالإضافة إلى معاناة معظم أدوات الكشف عن محتوى الذكاء الاصطناعي أوجهًا متعددة للقصور، لاسيما في خوارزمياتها ومستوى تغطية قواعدها، وكان من أبرز توصياتها ضرورة اتجاه الجامعات والمؤسسات الأكاديمية في مصر والعالم العربي بتطوير أداة للكشف عن محتوى الذكاء الاصطناعي داعمة للمحتوى العربي.

الكلمات المفتاحية: الذكاء الاصطناعي، أدوات الكشف عن نصوص الذكاء الاصطناعي، الانتحال العلمي، البيئة الأكاديمية.

Abstract:

The study aimed to introduce tools for detecting artificial intelligence content in terms of their general features, pricing systems, advantages and disadvantages, and to evaluate the extent of their effectiveness and accuracy in detecting plagiarism in the academic environment. To do this, the researcher relied on the descriptive analytical approach, and applied it to a sample whose vocabulary numbered 8 tools. , which in turn was used with several selected articles "human content, artificial intelligence content," numbering 4 articles specialized in the field of libraries and information, by measuring the distinguishing features of AI Detectors between articles written by artificial intelligence tools and articles written by humans, as well as measuring The accuracy of AI Detectors' recognition of AI articles subject to translation and AI articles subject to reformulation. One of its most important findings was that AI content detection tools with their current specifications do not appear to be ready enough to accurately and convincingly detect content generated by AI tools, in addition to Most tools for detecting artificial intelligence content suffer from multiple shortcomings, especially in their algorithms and the level of coverage of their rules. One of its most prominent recommendations was the necessity of directing universities and academic institutions in Egypt and the Arab world to develop a tool for detecting artificial intelligence content that supports Arabic content.

١- الإطار المنهجي للدراسة:

١/١- مقدمة:

حقق الذكاء الاصطناعي (AI) تقدمًا هائلاً في السنوات الأخيرة، بدءًا من النماذج التوليدية (GAI) وصولاً إلى نماذج اللغات الكبيرة (LLMs) في معالجة اللغة الطبيعية (Saharia, et.al, 2022)، حيث يمكن لهذه النماذج الآن إنشاء نصوص ذات جودة عالية مع إمكانية استخدامها في العديد من التطبيقات، كإنشاء نصوص شبيهة بالبشر مثل كتابة رموز البرمجة، وكلمات الأغاني، واستكمال المستندات، والإجابة على الأسئلة؛ .. ألخ من التطبيقات التي لا حصر لها. ويُظهر الاتجاه السائد في البرمجة اللغوية العصبية أن نماذج اللغات الكبيرة هذه سوف تتحسن بمرور الوقت (Sadasivan, et.al, 2023). ومن هنا يأتي التحدي الأكبر، حيث بات المجال مفتوحًا أمام إساءة الاستخدام، والتلاعب من قبل المستخدمين، وكذا انتشار الممارسات غير الأخلاقية، والتي شكلت في مجملها تحديات عظيمة أما المجتمع الأكاديمي الذي يسعى لضبط وتقنين الممارسات العلمية.

ويشير (Salvagno, Taccone & Gerli, 2023) إلى أنه على الرغم من أهمية أدوات الذكاء الاصطناعي، حيث يمكن لأداة مثل ChatGPT أن تصبح أداة قوية وواعدة لمهام كبيرة مثل إنشاء مسودات البحوث، وتلخيص المقالات، والترجمة، والتي قد تبدو مفيدة وقيمة في الأنشطة الأكاديمية لتسهيل عملية البحث والكتابة العلمية وتسريعها، إلا أن استخدام هذه الأداة في الكتابة العلمية يثير بعض المخاوف الأخلاقية التي تستوجب ضبط العلاقة مع هذه الأدوات وبخاصة في البيئة الأكاديمية.

وقد صاحب ظهور هذه التكنولوجيا العديد من التحديات، حيث يعتبرها البعض أكثر ذكاءً من البشر، ويعتبرها البعض الآخر تكنولوجيا مدمرة، وهناك من يعتبرها مساعدًا كبيرًا في الكتابة العلمية، إلا أن النقطة الأكثر خطورة هي استخدام بعض الباحثين الآن لهذه التكنولوجيا بوصفها مؤلف مشارك (King, M. R., &

ChatGPT, 2023)، مما يزيد من انتشار الممارسات غير الأخلاقية في البيئة الأكاديمية وبخاصة كتابة البحث العلمي.

كما شكل ظهور أدوات الذكاء الاصطناعي منعطفًا خطيرًا في إنتاج المعرفة، مما كان له تأثيرات مباشرة على كافة الأصعدة، فمع إطلاق ChatGPT في ٣٠ نوفمبر من عام ٢٠٢٢م أدى ذلك إلى التأثير على روبوتات الدردشة المدعومة بالذكاء الاصطناعي التي كانت قبل ChatGPT غير معروفة نسبيًا (Chaka, 2023)، كما أدت أيضًا إلى ظهور أدوات الكشف عن المحتوى المدعوم بالذكاء الاصطناعي AI Content Detector tools التي تهدف إلى الكشف والتمييز بين النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي والنصوص المكتوبة بواسطة البشر، ونظرًا لحداثة هذه الأدوات فإنها تعاني نقصًا في الدراسات التي تستكشف مدى فاعليتها وموثوقيتها في البيئة الأكاديمية.

وقد ظهرت أدوات الكشف عن المحتوى المدعوم بالذكاء الاصطناعي في الوقت الذي ظهرت فيه تأكيدات بأنه لا توجد أداة حالية لمحتوى الذكاء الاصطناعي يمكنها اكتشاف الانتحال العلمي للذكاء الاصطناعي في الاستجابات التي تم إنشاؤها بواسطتها (Anders, 2023; Caren, 2023)، في حين أنه يمكن استخدام هذه الأدوات لاكتشاف المحتوى المدعوم بالذكاء الاصطناعي في أنواع النصوص المختلفة، ذلك المحتوى الذي بات يشكل مصدرًا للإزعاج داخل المؤسسات الأكاديمية.

٢/١ – مشكلة الدراسة:

أكدت التطورات الحديثة في نماذج اللغات الكبيرة للمحولات التوليدية المدربة مسبقًا على المخاطر المحتملة للاستخدام غير الأخلاقي للذكاء الاصطناعي (AI) في البيئة الأكاديمية، وفي محاولات لمجابهة هذه المخاطر المحتملة، تم إصدار عدد من أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors التي تركز نشاطها في

اكتشاف المحتوى الناتج عن أدوات الذكاء الاصطناعي، وإنطلاقًا من هنا، تتبلور مشكة الدراسة في نقص التجارب الأكاديمية حول هذه الأدوات، وإمكاناتها وخصائصها، ولهذا سعت الدراسة الحالية إلى تقييم دقة وموثوقية هذه الأدوات للتعرف على النصوص المنشأة بواسطة أدوات الذكاء الاصطناعي، وكذا قدرتها على التمييز بين نصوص الذكاء الاصطناعي والنصوص المنشأة بواسطة البشر، بالإضافة إلى رصد مدى قدرتها في التعرف على النصوص المترجمة والنصوص المعاد صياغتها paraphrased بتطبيقها على مجموعة من المقالات المختارة في مجال علم المكتبات والمعلومات.

١/٣ - أهمية الدراسة:

ترجع أهمية الدراسة الحالية إلى أهمية موضوعها، الذي يمتاز بالحداثة على علم المكتبات والمعلومات، لاسيما في ظل انتشار أدوات الذكاء الاصطناعي، وتعدد وظائفها، بالشكل الذي باتت معه قادرة على صناعة محتوى أكاديمي يحاكي نظيره المعد بواسطة المؤلفين البشريين، مما يستلزم التوسع في أدوات الكشف عن المحتوى المدعوم بالذكاء الاصطناعي، يضاف إلى ذلك ما أشار إليه (chaka, 2023) من أن معظم أجهزة الكشف عن محتوى الذكاء الاصطناعي لا تزال جديدة، ولم يتم إجراء الكثير من الدراسات لتقييم مدى فعاليتها ودقتها وموثوقيتها من حيث التمييز بين المحتوى الذي تم إنشاؤه بواسطة روبوتات الدردشة التي تعمل بالذكاء الاصطناعي والمحتوى المكتوب بواسطة البشر، وعليه فإن هذا المجال جديد ومتنامي ولا يزال يحتاج إلى الكثير من الجهود البحثية، يضاف إلى ذلك الأهمية التطبيقية المتمثلة في تقييم موقف هذه الأدوات، ورصد فاعليتها لاسيما في البيئة العربية.

١/٤ - أهداف الدراسة:

١- التعرف على الملامح العامة لأدوات الكشف عن محتوى الذكاء الاصطناعي، ونظم تسعيرها، وكذا رصد إيجابياتها وسلبياتها.

٢- قياس ملامح تمييز أدوات الكشف AI Detectors بين المقالات المكتوبة بواسطة أدوات الذكاء الاصطناعي والمقالات المكتوبة بواسطة البشر.

٣- قياس دقة تعرف أدوات الكشف AI Detectors على مقالات الذكاء الاصطناعي الخاضعة للترجمة.

٤- رصد مستوى الدقة في تعرف أدوات الكشف AI Detectors على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة.

١/٥- تساؤلات الدراسة:

١- ما الملامح العامة لأدوات الكشف عن محتوى الذكاء الاصطناعي؟ وما نظم تسعيرها؟ وما أبرز إيجابياتها وسلبياتها؟

٢- ما مستوى تمييز أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors بين المقالات المكتوبة بواسطة أدوات الذكاء الاصطناعي والمقالات المكتوبة بواسطة البشر؟

٣- ما مدى الدقة في تعرف أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors على مقالات الذكاء الاصطناعي الخاضعة للترجمة.

٤- ما مستوى الدقة في تعرف أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة.

١/٦ - حدود الدراسة:

١/٦/٢- الحدود اللغوية: اعتمدت الدراسة على الأدوات الداعمة للعديد من اللغات، وجميعها لا يدعم العربية.

١/٦/٣- الحدود النوعية: شكلت أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors المحور النوعي للدراسة التحليلية.

١/٦/٤- الحدود الزمنية: امتدت الدراسة لتشمل أدوات الكشف عن محتوى الذكاء الاصطناعي المتاحة افتراضيًا حتى نهاية نوفمبر ٢٠٢٣م.

٧/١ – منهج الدراسة:

تنتمي الدراسة إلى الدراسات الاستكشافية، والتي يسعى من خلالها الباحث لاستكشاف موضوع أو مجال لم تتم دراسته من قبل وبخاصة في مجال المكتبات والمعلومات، ولا يزال تقييم فاعلية ودقة أدوات الكشف عن محتوى الذكاء الاصطناعي في التمييز بين المحتوى الذي ينشئه الذكاء الاصطناعي والمحتوى البشري مجالًا لم يحظ بدراسات أكاديمية موسعة في البيئة العربية، يضاف إلى ذلك اعتماد الدراسة على أسلوب الوصف والتحليل، وذلك بهدف التعرف على أدوات الكشف عن محتوى الذكاء الاصطناعي، ورصد مدى دقتها في اكتشاف النصوص العلمية بالتطبيق على مجال المكتبات والمعلومات، وذلك في أربعة محاور رئيسة:

- نصوص متخصصة تم إنشاؤها بواسطة أدوات الذكاء الاصطناعي.
- نصوص متخصصة تم إنشاؤها بواسطة مؤلفين بشريين.
- نصوص متخصصة تم إنشاؤها وترجمتها بواسطة أدوات الذكاء الاصطناعي.
- نصوص متخصصة تم إنشاؤها وإعادة صياغتها بواسطة أدوات الذكاء الاصطناعي.

حيث تم إخضاع هذه النصوص للفحص من خلال أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors، ومن ثم تحليل استجاباتها للتعرف على مستوى دقة هذه الأدوات في التعرف على محتوى الذكاء الاصطناعي.

٨/١ – مجتمع الدراسة وعينتها:

تشكل مجتمع الدراسة في أدوات الكشف عن محتوى الذكاء الاصطناعي AI Detectors، والتي أخذت في الانتشار لمجابهة أخطار الذكاء الاصطناعي في البيئة الأكاديمية، وقد تعددت هذه الأدوات بما تجاوز العشرين أداة، وهي تلك الأداوت التي أمكن للباحث التوصل إليها – في حدود علمه – من خلال شبكات التواصل الاجتماعي، ومحركات البحث، وأدوات الذكاء الاصطناعي، وقد وضع الباحث

مجموعة من المعايير لاختيار عينة من هذه الأدوات وهي (إتاحة الحسابات المجانية – دعم العديد من اللغات – ثبات الإتاحة)، وبعد إخضاع هذه الأدوات لهذه المعايير الثلاث، وأسفر ذلك عن تحقق هذه المعايير في ثماني أدوات، والتي شكلت العينة النهائية للدراسة، وذلك بنسبة ٤٠% من مجتمع الدراسة، والتي ينقلها الجدول (١) تفصيلاً.

جدول (١) عينة الدراسة

١	Copyleaks	https://copyleaks.com/ai-content-detector
٢	Content at scale	/https://contentatscale.ai/ai-content-detector
٣	Undetectable AI	/ https://undetectable.ai
٤	GptZero	/ https://gptzero.me
٥	Crossplag	/https://crossplag.com/ai-content-detector
٦	Writer AI	/https://writer.com/ai-content-detector
٧	Sapling	https://sapling.ai/ai-content-detector
٨	Winston AI	https://app.gowinston.ai/projects/1338034

١/٩- مصطلحات الدراسة (تعريفات إجرائية):

تزخر الدراسة بعدد من المصطلحات، والتي يعد من أبرزها ما يلي:

- **أدوات الكشف عن محتوى الذكاء الاصطناعي:** الأدوات التي تركز نشاطها في اكتشاف المحتوى الناتج عن أدوات الذكاء الاصطناعي، كالكتابة، وإعادة الصياغة، والترجمة..الخ.
- **الذكاء الاصطناعي التوليدي:** تقنية يمكنها إنتاج محتوى جديد بشكل تلقائي وفعال عن طريق استخدام الخوارزميات والنماذج الذكية لإنشاء محتوى جديد، سواء كانت نصوص، أو صور، أو مقاطع صوتية، أو مقاطع فيديو..الخ.

- **الانتحال العلمي:** يقصد بها تقديم أعمال علمية بشكل غير أمين اعتمادًا على أدوات الذكاء الاصطناعي، والتي قدد تتعدد أوجهة بين التأليف العلمي، والترجمة، وإعادة صياغة النصوص.

١٠/١- الدراسات السابقة:

بعد مراجعة الباحث للإنتاج الفكري المتخصص حول موضوع الدراسة الحالية، واعتماده على العديد من قواعد البيانات العالمية " ProQuest, Emerald, Science Direct, Scopus, Google scholar Academic Search Complete," بالإضافة إلى قواعد بيانات دار المنظومة، لم يتوصل الباحث لدراسات عربية حول "أدوات الكشف عن محتوى الذكاء الاصطناعي" موضوع الدراسة، إلا أنه أمكن التوصل لمجموعة من الدراسات الأجنبية الحديثة التي تناولت جانبًا أو أكثر من جوانب الموضوع، ويعرض الباحث لها فيما يلي مرتبة ترتيبًا هجائيًا حسب اسم المؤلف لأن جميعها يرجع لعام ٢٠٢٣م.

دراسة (Akram, 2023) التي هدفت إلى تحليل ستة أدوات لتعريف النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي من خلال تقييم نقاط القوة والقيود الخاصة بها عبر العديد من المعايير، وذلك من خلال إنشاء مجموعة بيانات متعددة المجالات لاختبار أحدث واجهات برمجة التطبيقات وأدوات اكتشاف المعلومات المولدة والتي تستخدمها الجامعات والمؤسسات البحثية الأخرى. وتم إنشاء مجموعة بيانات كبيرة تتكون من المقالات والملخصات والقصص والأخبار ومراجعات المنتجات، ومن ثم تم استخدام مجموعة البيانات التي تم إنشاؤها حديثًا لوضع ستة أدوات في خطواتها، وهي: "GPTkit" و"GPTZero" و"Originality" و"Sapling"و"Writer" "Zylalalab"،

وتشير النتائج إلى تراوح معدلات دقة هذه الأدوات بين ٥٥.٢٩ و٩٧.٠%، وعلى الرغم من أن جميع الأدوات حققت نجاحًا جيدًا في التقييمات، إلا أن الأصالة

كانت فعالة بشكل خاص في جميع المجالات، وقد أوصت الدراسة بالحاجة إلى مزيد من الدراسة لمعرفة مدى فعاليتها بالنسبة لمواد ChatGPT متعددة المجالات.

دراسة (Cingillioglu, 2023) التي هدفت إلى دراسة الاستراتيجيات اللازمة لمواجهة تحدي الحفاظ على النزاهة الأكاديمية في البيئات التعليمية، من خلال تعريف نموذج لغوي لتحقيق دقة عالية في التمييز بين المقالات التي تم إنشاؤها بواسطة ChatGPT والمقالات المكتوبة بواسطة الإنسان مع التركيز بشكل خاص على تصنيف المقالات المكتوبة بواسطة الإنسان "ليس بشكل خاطئ" على أنها مقالات تم إنشاؤها بواسطة الذكاء الاصطناعي، وقد كشفت الدراسة عن تسجيل دقة ١٠٠% من خلال خوارزمية آلة ناقل الدعم (SVM) لتحديد المقالات التي أنشأها الإنسان، كما ناقش المؤلف الاستخدام الرئيسي لنقاط الاستدعاء ودرجة F2 لقياس أداء التصنيف وأهمية التخلص من السلبيات الكاذبة والتأكد من عدم تصنيف أي مقالات فعلية أنشأها الإنسان بشكل غير صحيح على أنها مقالات تم إنشاؤها بواسطة الذكاء الاصطناعي. وتمت مقارنة نتائج خوارزميات تصنيف النموذج المقترح مع نتائج برامج الكشف عن النصوص المولدة بواسطة الذكاء الاصطناعي والتي طورتها OpenAI وGPTZero وCopyleaks، وقد أوصت الدراسة بإمكانية اعتماد المدرسين والمصممين التعليميين اكتشاف المقالات التي أنشأها الطلاب بواسطة الذكاء الاصطناعي باستخدام نموذج اللغة المقترح ومصنف التعلم الآلي (ML) بدقة عالية، بل ويجب، تحديد المقالات التي يكتبها الطلاب بشكل صحيح وبدقة ١٠٠% حتى لو انخفض أداء دقة التصنيف الإجمالي قليلاً.

دراسة (Elkhatat, Elsaid, & Almeer, 2023) التي هدفت إلى تحليل قدرات أدوات الكشف عن محتوى الذكاء الاصطناعي المختلفة في تمييز المحتوى البشري والمحتوى المؤلف بواسطة الذكاء الاصطناعي، حيث تم إنشاء خمسة عشر فقرة من نموذجي ChatGPT 3.5 و٤ حول موضوع أبراج التبريد في العملية الهندسية

وخمس استجابات تحكم من صنع الإنسان للتقييم. وتم استخدام أدوات الكشف عن محتوى الذكاء الاصطناعي التي طورتها شركات OpenAI وWriter وCopyleaks وGPTZero وCrossPlag لتقييم هذه الفقرات. تكشف النتائج أن أدوات اكتشاف الذكاء الاصطناعي كانت أكثر دقة في تحديد المحتوى الناتج عن GPT 3.5 مقارنة بـ GPT 4. ومع ذلك، عند تطبيقها على استجابات التحكم المكتوبة بواسطة الإنسان، أظهرت الأدوات تناقضات، مما أدى إلى نتائج إيجابية كاذبة وتصنيفات غير مؤكدة. وهذا ما دفع هذه الدراسة إلى أن توص بضرورة الحاجة إلى مزيد من التطوير والتحسين لأدوات الكشف عن محتوى الذكاء الاصطناعي حيث يصبح المحتوى الناتج عن الذكاء الاصطناعي أكثر تعقيدًا ويصعب تمييزه عن النص المكتوب بواسطة الإنسان.

دراسة (Chaka, 2023) التي هدفت إلى اختبار دقة خمس أدوات لمحتوى الذكاء الاصطناعي، وهي: GPTZero، OpenAI Text Writer.com's AI Content Detector, Copyleaks AI Content Detector, and Giant Language model Test Room، وذلك لاكتشاف المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي في الاستجابات التي تم إنشاؤها بواسطة ChatGPT وYouChat وChatsonic. حيث تم إنشاء الاستجابات من روبوتات الدردشة الثلاثة هذه باستخدام الذكاء الاصطناعي باستخدام المطالبات الإنجليزية المتعلقة بدراسات اللغة الإنجليزية التطبيقية. بعد ذلك، تمت ترجمة الاستجابات التي تم إنشاؤها بواسطة ChatGPT بواسطة Google إلى الألمانية والفرنسية والإسبانية وجنوب سوتو وإيزولو، وإدخالها في GPTZero لاكتشاف المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي فيها. بالإضافة إلى ذلك، تم إدخال الردود التي تم إنشاؤها بواسطة ChatGPT والمترجمة من Google إلى الألمانية والفرنسية والإسبانية في Copyleaks AI Content Detector حتى يتمكن من اكتشاف المحتوى الذي تم

إنشاؤه بواسطة الذكاء الاصطناعي فيها. بالنسبة للاستجابات التي تم إنشاؤها بواسطة ChatGPT وYouChat وChatsonic، برز Copyleaks AI Content Detector باعتباره كاشف محتوى الذكاء الاصطناعي الأفضل أداءً بين أجهزة الكشف عن محتوى الذكاء الاصطناعي الخمسة. تبعه OpenAI Text. فيما يتعلق بالاستجابات التي تم إنشاؤها بواسطة ChatGPT والتي تمت ترجمتها بواسطة Google إلى خمس لغات، فقد أخطأ GPTZero في تعريفها جميعًا على أنها من إنتاج الإنسان. بالنسبة للاستجابات التي تم إنشاؤها بواسطة ChatGPT والتي تمت ترجمتها بواسطة Google إلى الألمانية والفرنسية والإسبانية، حدد كاشف محتوى Copyleaks AI بشكل صحيح ثلاثة من النصوص المترجمة إلى الألمانية، وخمسة من النصوص المترجمة إلى الفرنسية، وجميع النصوص المترجمة إلى الإسبانية على أنها بواسطة الذكاء الاصطناعي، وكان من أهم نتائجها أن جميع أجهزة الكشف عن محتوى الذكاء الاصطناعي الخمسة تبدو غير جاهزة تمامًا بعد للكشف الدقيق والمقنع عن المحتوى الناتج عن الذكاء الاصطناعي من النصوص المولدة آليًا في سياقات مختلفة. وهذا له عواقب وخيمة على الانتحال الناتج عن الذكاء الاصطناعي في كتابة المقالات الأكاديمية.

دراسة (Sadasivan, et,al, 2023) التي هدفت إلى التعرف على إمكانية اكتشاف النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي بشكل موثوق بواسطة أدوات الكشف عن الذكاء الاصطناعي، وذلك في ظل الاستخدام غير المنظم لنماذج اللغات الكبيرة، حيث يمكن أن يؤدي إلى عواقب ضارة مثل الانتحال، وتوليد أخبار مزيفة، والبريد العشوائي، وما إلى ذلك، حيث عكست الدراسة أن هذه الكاشفات ليست موثوقة في السيناريوهات العملية، ومن الناحية التجريبية، أظهرت أن هجمات إعادة الصياغة، يمكن أن تخدع مجموعة كاملة من أجهزة الكشف، بما في ذلك تلك التي تستخدم مخططات العلامات المائية بالإضافة إلى أجهزة الكشف القائمة على الشبكة

العصبية ومصنفات الطلقة الصفرية، كما تبين أن النصوص المنشأة بواسطة نماذج اللغات الكبيرة المحمية بواسطة مخططات العلامات المائية يمكن أن تكون عرضة لهجمات الانتحال، حيث يمكن للبشر المنافسين استنتاج التوقيعات النصية المخفية LLM وإضافتها إلى النص الذي أنشأه الإنسان ليتم اكتشافه كنص تم إنشاؤه بواسطة اللغات الكبيرة، مما قد يتسبب في الإضرار بسمعة المطورين، وثقد أوصت الداسة بضرورة فتح محادثة صادقة مع المجتمع فيما يتعلق بالاستخدام الأخلاقي والموثوق للنص الذي ينشئه الذكاء الاصطناعي.

دراسة (Walters, 2023) التي هدفت إلى تقييم دقة ١٦ جهازًا للكشف عن نص الذكاء الاصطناعي في التمييز بين الكتابة التي ينشئها الذكاء الاصطناعي والكتابة التي ينشئها الإنسان، حيث تضمنت المستندات التي تم تقييمها ٤٢ مقالة جامعية تم إنشاؤها بواسطة ChatGPT-3.5، و٤٢ مقالة تم إنشاؤها بواسطة ChatGPT-4، و٤٢ مقالة كتبها الطلاب في دورة تكوينية في السنة الأولى دون استخدام الذكاء الاصطناعي، حيث تم تقييم أداء كل كاشف فيما يتعلق بدقته الإجمالية، ودقته مع كل نوع من المستندات، وحسمه (العدد النسبي للاستجابات غير المؤكدة)، وعدد الإيجابيات الخاطئة (الأوراق التي أنشأها الإنسان والتي تم تحديدها على أنها ذكاء اصطناعي بواسطة الكاشف)، وعدد السلبيات الكاذبة (الأوراق التي تم إنشاؤها بواسطة الذكاء الاصطناعي والمصنفة على أنها بشرية)، وقد أظهرت النتائج تمتع أجهزة الكشف الثلاثة – Copyleaks، وTurnItIn، وOriginality.ai بدقة عالية في التعامل مع المجموعات الثلاث من المستندات، وعلى الرغم من أن معظم الكاشفات الـ ١٣ الأخرى يمكنها التمييز بين أوراق GPT-3.5 والأوراق التي أنشأها الإنسان بدقة عالية إلى حد معقول، إلا أنها غير فعالة بشكل عام في التمييز بين أوراق GPT-4 وتلك التي كتبها طلاب المرحلة الجامعية. وبشكل عام، فإن أجهزة الكشف التي تتطلب التسجيل والدفع تكون أكثر دقة قليلاً من أجهزة الكشف الأخرى.

دراسة (Weber-Wulff, et.al, 2023) التي سعت إلى التعرف على أدوات الكشف عن الذكاء الاصطناعي، التي يمكنها التمييز بشكل موثوق بين النص المكتوب بواسطة الإنسان والنص الذي تم إنشاؤه بواسطة ChatGPT، وما إذا كانت تقنيات الترجمة الآلية وتشويش المحتوى تؤثر على اكتشاف النص الذي تم إنشاؤه بواسطة الذكاء الاصطناعي من عدمه، حيث تكونت عينة الدراسة من ١٢ أداة متاحة للعامة ونظامين تجاريين (Turnitin وPlagiarismCheck) يستخدمان على نطاق واسع في البيئة الأكاديمية، وقد خلصت الدراسة إلى أن أدوات الكشف المتاحة ليست دقيقة ولا موثوقة ولديها تحيز رئيس نحو تصنيف المخرجات على أنها مكتوبة بواسطة الإنسان بدلاً من اكتشاف النص الذي تم إنشاؤه بواسطة الذكاء الاصطناعي. علاوة على ذلك، تؤدي تقنيات تشويش المحتوى إلى تفاقم أداء الأدوات بشكل كبير. وقد قدمت الدراسة العديد من المساهمات المهمة مثل: تلخصها أحدث الجهود العلمية وغير العلمية المماثلة في هذا المجال، عرضها نتيجة أحد الاختبارات الأكثر شمولاً التي تم إجراؤها حتى الآن، مناقشة الآثار المترتبة وعيوب استخدام أدوات الكشف عن النص الناتج عن الذكاء الاصطناعي في الأوساط الأكاديمية.

دراسة (Xiang, Zhou, & Xie, 2023) التي هدفت إلى رصد ثقة المستخدمين في أدوات الذكاء الاصطناعي واعتمادهم على نتائجها للكشف عن مراجعات البريد العشوائي واستكشاف تأثير العمر، من خلال التلاعب بمعايير الكشف عن نصوص الذكاء الاطصناعي وكذا نتائجه، حيث تم تصميم نوعين من الشرح، أداة Text-AI وأداة Behaviour-AI. تحدد أداة Text-AI معايير الكشف بناءً على السمات النصية للمراجعات، بينما تحدد أداة Behaviour-AI معايير الكشف بناءً على السمات السلوكية للمراجعين. تم قياس ثقة البالغين الأصغر سنًا (٢٠–٢٦ عامًا) وكبار السن (٥٠–٧٨ عامًا) في أدوات الذكاء الاصطناعي والتغيرات في أحكام المصداقية للمراجعات والموقف العام تجاه المنتج الذي قدموه بناءً على نتائج اكتشاف

أدوات الذكاء الاصطناعي، وقد أظهرت نتائج الدراسة ثقة جميع المشاركين الأكبر سنًا تقريبًا في تنبؤات أدوات الذكاء الاصطناعي، لكن ٤٨.٧% منهم قد يتخلون عن الذكاء الاصطناعي عندما يكون تنبؤ الذكاء الاصطناعي مختلفًا عن حكمهم الخاص، كما أظهر البالغون الأصغر سنًا ثقة أعلى في أداة Behaviour-AI مقارنة بأداة Text-AI، خاصة عندما اكتشفت أداة الذكاء الاصطناعي عددًا أكبر من مراجعات البريد العشوائي أكثر من أنفسهم، كما رأى المشاركون أن أدوات الذكاء الاصطناعي أكثر كفاءة وإحسانًا عندما تفوقت أدوات الذكاء الاصطناعي على المشاركين من خلال الكشف عن مراجعات البريد العشوائي غير المتوقعة، وقد وقد أوصت الدراسة بضرورة تحمل وسائل الإعلام ومنصات التسوق مسؤولية جعل كبار السن أكثر وعيًا بوجود المراجعات المزيفة ومرسلي البريد العشوائي عبر الإنترنت، والضرر الذي قد ينجم عن ذلك، بالإضافة إلى ضرورة أخذ المعرفة الضعيفة لدى كبار السن حول مراجعات البريد العشوائي عند تصميم أدوات الذكاء الاصطناعي لهذه الفئة الضعيفة.

١/١٠/١- التعليق على الدراسات السابقة:

بتحليل الدراسات السابقة يتبين تركيز جلها على أدوات الكشف عن محتوى الذكاء الاصطناعي بصفة عامة، فهناك دراسات ركزت على تحليل أدوات تعريف النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي من خلال تقييم نقاط القوة، ومواطن الضعف، والقيود الخاصة بها عبر العديد من المعايير، وهناك دراسات حللت الاستراتيجيات اللازمة لمواجهة تحدي الحفاظ على النزاهة الأكاديمية في البيئات التعليمية، وهناك من ركز على تحليل قدرات أدوات الكشف عن محتوى الذكاء الاصطناعي المختلفة في تمييز المحتوى البشري والمحتوى المؤلف بواسطة الذكاء الاصطناعي، واختبار دقتها في اكتشاف النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي بشكل موثوق بواسطة أدوات الكشف عن الذكاء الاصطناعي، وهناك دراسات ركزت على تقييم دقة أدوات الكشف عن نص الذكاء الاصطناعي في التمييز

بين الكتابة التي ينشئها الذكاء الاصطناعي والكتابة التي ينشئها الإنسان، وأخرى رصدت ثقة المستخدمين في أدوات الذكاء الاصطناعي واعتمادهم لنتائجها للكشف عن مراجعات البريد العشوائي. يضاف إلى ذلك تعدد البنية المنهجية لهذه الدراسات، وإن كان الأسلوب الوصفي التحليلي أكثرها توظيفًا، كما تشابهت عدد من الدراسات في عيناتها من أدوات الكشف عن محتوى الذكاء الاصطناعي، وتتميز الدراسة الحالية عن تلك الدراسات بأنها أول دراسة عربية تهدف لاسكشاف إمكانات هذه الأدوات، ورصد معدلات دقتها بالتطبيق على نصوص متخصصة في مجال المكتبات والمعلومات، يضاف إلى ذلك تميزها في مجمل أهدافها التي جاءت أكثر تخصيصًا وتركيزًا على رصد معدلات دقتها بالتطبيق على نصوص متخصصة، فضلاً عن معيارية اختيار عينة أدوات الكشف، وعلى الرغم من اختلاف زوايا التناول بين الدراسة الحالية وبعض الدراسات السابقة إلا أنها شكلت في مجملها إضاءات نظرية ومنهجية أفادت الدراسة الحالية.

٢- الإطار النظري للدراسة:

٣- ٢/١- أدوات الذكاء الاصطناعي والضوابط الأخلاقية:

تشير بعض الدراسات إلى استحالة اكتشاف النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي، لاسيما مع تقدم النماذج اللغوية، حيث تقدمت قدرتها على محاكاة النص البشري، في الواقع، تصبح مشكلة اكتشاف النص باستخدام الذكاء الاصطناعي أكثر أهمية وإثارة للاهتمام في ظل وجود نماذج لغوية مصممة لتقليد البشر والتهرب من الاكتشاف، ومع التقدم في نماذج اللغة الكبيرة LLMs، أصبح توزيع النصوص التي أنشأها الذكاء الاصطناعي مشابهًا بشكل متزايد للنصوص التي أنشأها الإنسان، مما يجعل اكتشافها أكثر صعوبة (Lu, et.al, 2023).

تكمن معضلة الذكاء الاصطناعي في التعليم والبحث في أن المحتويات المولدة آليًا Machine Generated Contents (MGC) تدور حول التوازن بين الكفاءة

والأصالة. من ناحية، حيث يمكن لأدوات الذكاء الاصطناعي إنشاء المحتوى بكميات كبيرة وأنواع متعددة مع توفير عنصري السرعة والدقة إلى حد بعيد، والتي يمكن أن تكون مفيدة للأغراض التعليمية. ومن ناحية أخرى، هناك خطر التضحية بالتفرد والإبداع الذي يأتي من الكتابة الإنسانية وفن الإنتاج الفكري (Dalalah & Dalalah, 2023). ونتيجة لذلك، ومن أجل ضمان تعظيم مزايا الذكاء الاصطناعي مع تجنب عيوبه المحتملة، يجب أن يكون البحث والتطوير مستمرين في هذا المجال مصحوبًا بمعايير وضوابط أخلاقية.

وفي الوقت الذي ينقسم فيه المجتمع الأكاديمي بين مؤيد ومعارض، فهناك من يعتبر هذه التكنولوجيا مدمرة للإبداع وغير أخلاقية، وهناك من يعتبرها أداة ذكية ومفيدة وداعمة للإبداع، واقترح (Mohammadzadeh, et.al, 2023) طريقتان يمكنهما مساعدة المجتمعات الأكاديمية على منع الانتحال الأكاديمي: أولاً، حظر أدوات الذكاء الاصطناعي وتعديل أساليب التدريس، فقد حظرت العديد من الجامعات والمدارس استخدام ChatGPT داخلها، كما قامت مؤسسات أخرى بحظر الوصول إليها من شبكات جامعاتها، كما حظرت بعض المؤتمرات بشكل صريح المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي في عروض المؤتمرات، بما في ذلك مؤتمرات التعلم الآلي (Weber-Wulff, et.al, 2023) كما قامت بعض البلدان بحظر استخدام هذه الأدوات، كما هو الحال في إيطاليا على الرغم من إلغاء هذا القرار حاليًا (Schechner 2023) كما قام العديد منا الناشرين بحظر الاعتراف بأي أداة ذكاء اصطناعي قائمة على نماذج لغوية كبيرة كمؤلفين في دورياتهم (Nature, 2023 & Thorp, 2023) وثانيًا، الكشف عن المحتويات المنتجة بواسطة الذكاء الاصطناعي، فإلى جانب أدوات إنشاء المحتوى بواسطة الذكاء الاصطناعي، أصبحت أدوات التعرف على المحتوى الذي ينتجه الذكاء الاصطناعي مزدهرة أيضًا، والتي يمكنها تحديد النص وتصنيفه على أنه نص تم إنشاؤه بواسطة الإنسان أو بواسطة الآلة.

٢/٢- الذكاء الاصطناعي وخطر الانتحال العلمي:

يمكن تقسيم استخدام الذكاء الاصطناعي (AI) في الكتابة الأكاديمية إلى فئتين عريضتين: تلك التي تساعد المؤلفين في عملية الكتابة؛ وتلك التي تستخدم لتقييم جودة وصحة العمل المكتوب، فأدوات مثل معالجة اللغة الطبيعية التي يمكنها فهم وتوليد لغة شبيهة بالإنسان يمكن أن تساعد المؤلفين في كتابة الأوراق العلمية وإعدادها، كما يمكن لأدوات مثل برامج كشف الانتحال ومنصات مراجعة النظراء الآلية أن تساعد المراجعين والمحررين في تقييم جودة العمل العلمي (Golan, et.al, 2023).

ويشكل الانتحال عند استخدام أدوات الذكاء الاصطناعي للكتابة العلمية مصدر قلق محتمل يجب أن يؤخذ بعين الاعتبار، حيث يمكن لأدوات الذكاء الاصطناعي إنشاء نصوص تشبه نصوصًا من مصادر أخرى وبخاصة المعدة بشريًا، وبالتالي قد تنتج نصًا يمكن تمييزه على أنه انتحال، يضاف إلى ذلك خطر الاعتماد الكلي على هذه الأدوات مما يجعل النصوص المنتجة انتحالاً علميًا من أدوات الذكاء الاصطناعي.

ففي الوقت الذي يمكن لأدوات الذكاء الاصطناعي أن تكون فاعلة في عملية الكتابة العلمية، إلا أنه من المهم ملاحظة أنها ليست مثالية، وقد تمكنها الوقوع في أخطاء علمية أو تقديم معلومات غير دقية، وعليه فإن الإشراف البشري ضروري لضمان أن المحتوى الذي تم إنشاؤه يتمتع بقدر من الدقة، والموضوعية، ومناسبة الجمهور المستهدف (Huang & Tan, 2023).

وعلى الرغم من أن استخدام أدوات الذكاء الاصطناعي قد يكون مفيدًا في تسريع عملية الكتابة العلمية وتوليد الأفكار، فإنه من الأهمية بمكان تجنب مخاطر الانتحال عند استخدام نص تم إنشاؤه بواسطة الذكاء الاصطناعي، وقد يكون ذلك بالسبل الآتية (Huang & Tan, 2023):

- **فهم مصدر النص:** فمن المهم فهم مصدر النص، حيث تستخدم أدوات الذكاء الصطناعي مجموعة بيانات كبيرة لإنشاء نص جديد، لذلك هناك خطر من أن النص الذي تم إنشاؤه قد يتضمن محتوى تم نشره مسبقًا، ومن المهم مراجعة النص الذي تم إنشاؤه بعناية للتأكد من أنه ليس مجرد نسخة معاد صياغتها من المحتوى المنشور مسبقًا.
- **استخدام مصادر متعددة:** فمن المهم استخدام مصادر متعددة عند كتابة مقال علمي، لتجنب خطر الانتحال، حيث يمكن استخدام هذه الأدوات لإنشاء نص بناءً على مصدر واحد أو أكثر، ولكن من المهم استكمال النص المُنشأ بمعلومات من مصادر أخرى، مما سيساعد في التأكد من أن مقالة المراجعة ليست مجرد تكرار للمحتوى المنشور مسبقًا.
- **الاستشهاد بالمصادر بشكل صحيح:** حيث يمكن إنشاء استشهادات ومراجع بناءً على نص الإدخال، ومع ذلك قد تقترح أدوات الذكاء الاصطناعي مراجع خاطئة أو حتى الاستشهاد بمقالات غير موجودة؛ لذلك من المهم للغاية مراجعة الاستشهادات والمراجع التي تم إنشاؤها وتحريرها لضمان دقتها واكتمالها.
- **استخدام برامج الكشف عن الانتحال:** حيث يمكن استخدام برامج الكشف عن الانتحال لتحديد حالات الانتحال في المقالة العلمية، مما يساعد في التأكد من أن المقالة أصلية وليست مجرد نسخة معاد صياغتها من المحتوى المنشور مسبقًا.
- **المراجعة والتحرير بعناية:** فمن المهم مراجعة المقالة العلمية وتعديلها بعناية للتأكد من أنها أصلية ولا تتضمن أي حالات سرقة أدبية، ويتضمن ذلك مراجعة النص الذي تم إنشاؤه أو تحريره لضمان أصالته ودقته.

٣/٢ – ممارسات كشف محتوى الذكاء الاصطناعي:

مع التقدم المستمر للذكاء الاصطناعي التوليدي، واتساع دائرة انتشاره، وتعدد أدواته، وما صحبه من تطويرات في معالجة اللغة الطبيعية، باتت هذه الأدوات تمتلك

قدرات استثنائية في إنتاج نصوص مترابطة وملائمة سياقيًا، اكتسب تطوير منصات الكشف عن محتوى الذكاء الاصطناعي زخمًا كبيرًا، مما أثار مناقشات علمية بشأن صحة المحتوى الذي تم إنشاؤه آليًا وتدابير اكتشافه.

وتشير الدراسات إلى تطوير مجتمع البحث مجموعة واسعة من أساليب الكشف عن المحتويات التي يولدها الذكاء الاصطناعي، والتي يمكن تصنيفها في ثلاث فئات رئيسة:

١- **الأساليب القائمة على التدريب:** تتعامل الأساليب القائمة على التدريب مع مشكلة الكشف باعتبارها مهمة تصنيف ثنائية، حيث يتم تدريب الشبكات العصبية باستخدام النص الذي تم إنشاؤه بواسطة الذكاء الاصطناعي والنص المكتوب بواسطة الإنسان، وقد استخدمت الدراسات المبكرة المصنفات لتحديد المراجعات المزيفة، والأخبار المزيفة، كما قام الباحثون مؤخرًا بتدريب المصنفات باستخدام النصوص التي تم إنشاؤها بواسطة نماذج اللغات الكبيرة (Guo, ET.AL, 2023).

٢- **الأساليب الإحصائية:** تركز الأساليب الإحصائية على اكتشاف الطلقة الصفرية دون أي تدريب إضافي، حيث يسعى هذا الأسلوب إلى التمييز بين النص المكتوب بواسطة الإنسان والنص الذي ينشئه الذكاء الاصطناعي استنادًا إلى الخصائص الإحصائية للنص، مثل المخالفات الإحصائية في مقاييس مثل الإنتروبيا، والرتبة الرمزية..الخ (Mitchell,ET.AL, 2023)، وتستغل الطريقة الحديثة الظاهرة المتمثلة في أن النص الناتج عن الذكاء الاصطناعي يميل إلى التواجد في مناطق الانحناء السلبي لسجل احتمالية النص.

٣- **العلامات المائية:** تتضمن طرق العلامة المائية تعديل عملية إنشاء النص في نماذج اللغات الكبيرة لطباعة أنماط محددة على النص الذي تم إنشاؤه، بحيث يمكن اكتشافه (Kirchenbauer, ET.AL, 2023).

٣- نتائج الدراسة ومناقشتها:

١/٣ - حجم الحروف المتاحة للفحص بأدوات الدراسة:

ينقل الجدول (٢) حجم الحروف المتاحة للفحص بأدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة.

جدول (٢) حجم الحروف المتاحة للفحص بأدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة

الأداة	الحد الأدني للحروف	الحد الأقصى للحروف والكلمات
Copyleaks	٣٥٠ حرف	غير محدد
Content at scale	٢٥ حرف	٢٥٠٠ حرف
Undetectable AI	٥٠ حرف	غير محدد
GptZero	٢٥٠ حرف	٥٠٠٠ حرف
Crossplag	غير محدد	٣٠٠٠ كلمة
Writer AI	غير محدد	١٥٠٠ حرف
Sapling	٥٠ حرف	غير محدد
Winston AI	٦٠٠ حرف	٢٠٠٠ كلمة

بتحليل بيانات الجدول (٢) يتبين التفاوت النسبي بين أدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة في الحدود الدنيا والقصوى للأحرف المسموح بتحليلها من قبل هذه الأدوات، فمن حيث تحديد الحروف من عدمه بلغت نسبة الأدوات المحددة للحد الأدنى للحروف ٧٥% مقابل نسبة ٢٥% من الأدوات لم تحدد هذا الحد الأدنى، أما من حيث الحد الأقصى لعدد الحروف فبلغت نسبة الأدوات المحددة لهذا الحد ٦٢.٥% مقابل نسبة ٣٧.٥% من الأداوت لم تحدد هذا الحد الأقصى.

كما جاءت أداة Winston AI أفضلها على الإطلاق في تحديد الحدود الدنيا للحروف، وذلك بواقع ٦٠٠ حرفًا كحد أني للتحليل، تلتها أداة Copyleaks والتي حددت الحدود الدنيا بواقع ٣٥٠ حرفًا، كما كانت أدناها على الإطلاق أداة content at scale وذلك بواقع ٢٥ حرفًا فقط، في حين لم تحدد منصتا: Crossplag,

Writer AI أي حدود دنيا للتحليل بمنصاتها، وفيما يتعلق بالحدود القصوى، فقد احتلت أداة Crossplag مركز الصدارة، وذلك بواقع تحليلها ٣٠٠٠ كلمة كحد أقصى للتحليل، والذي يعد نسبة كبيرة بين تلك الأدوات، تلتها في ذلك أداة Winston AI وذلك بواقع ٢٠٠٠ كلمة كحد أقصى للتحليل، كما جاءت أداة Writer AI أدناها على الإطلاق، بواقع ١٥٠٠ حرف كحد أقصى للتحليل، بينما لم تحدد أدوات: Copyleaks, Undetectable AI, sapling أية حدود قصوى للتحليل، وقد كان التسجيل شرطًا رئيسًا في هذه الأدوات، والتي كانت تقبل التسجيل بحساب جوجل.

ويرى الباحث أن هذه المؤشرات، لا تعبر بالضرورة عن مدى جودة الفحص من قبل هذه الأدوات، فربما تتفوق أداة ذات حدود دنيا للفحص على أخرى ذات حدود أكبر.

٢/٣ – الملامح العامة لأدوات الكشف عن محتوى الذكاء الاصطناعي:

تتعدد الخصائص الرئيسة لأدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة، كما تتفاوت هذه الأدوات في نظم التسعير، والتي جاءت معلنة على منصاتها، والتي ينقلها الجدول (٣) تفصيلاً.

جدول (٣) الملامح العامة لأدوات الكشف عن محتوى الذكاء الاصطناعي

الأداة	الخصائص الرئيسية	نظم التسعير
Copyleaks	أداة تقدم تقييم دقيق على مستوى الجملة النصية الذي يفحص مصداقية النص بدقة، يحصل المستخدمون على رؤى ليس فقط في تمييز المقاطع البشرية مقابل الاصطناعية ولكن أيضًا في تسليط الضوء على الأجزاء المحتملة للتشبيه لإخفاء توليد الذكاء الاصطناعي باستخدام أحدث التقنيات.	– لدى كاشف Copyleaks AI خيار مجاني. – تبدأ الاشتراكات من ٩.٩٩ دولارًا شهريًا، وتغطي ١٠٠ نقطة عبر حسابين للمستخدمين. مدقق الانتحال الخاص بهم هو ١٠.٩٩ دولارًا شهريًا. تبلغ تكلفة تجميع الذكاء الاصطناعي واكتشاف الانتحال في حل واحد ١٦.٩٩ دولارًا شهريًا.

الأداة	الخصائص الرئيسية	نظم التسعير
		– تعتمد مقاييس التسعير على الاستخدام، حيث يتيح كل رصيد مسح ٢٥٠ كلمة ضوئيًا - إنفاق أكثر من ٢ نقطة عندما تتجاوز ٢٥٠ كلمة.
Content at scale	هو أداة كاشفة مبنية على السحابة تستخدم تعلم الآلة لتحديد محتوى تم إنشاؤه بواسطة الذكاء الاصطناعي. تم تصميمها لاستخدامها من قبل الشركات لاكتشاف محتوى تم إنشاؤه بواسطة الذكاء الاصطناعي في مواد التسويق وتفاعلات خدمة العملاء وغيرها من أشكال الاتصال الشركي.	– يتم تقديم Content at Scale بتكلفة مجانية، مما يتيح للمستخدمين تحليل ما يصل إلى ٢٥٠٠ حرف من النص لكل تحليل. – توفر الطبقة المجانية وسيلة لاختبار الوظائف الأساسية لفحوصات الكود المصغرة دون التزام مالي. – هناك أيضًا طبقة مدفوعة بسعر ٤٩ دولارًا شهريًا.
Undetectable AI	يقدم الكشف بواسطة الذكاء الاصطناعي وتنقيح المحتوى في حل واحد، حيث يدعم ١٠,٠٠٠ حرف في التحليل الواحد. باستخدام آخر التطورات في نماذج اللغة، كما إنه يعترف بالنص بشكل موثوق من خلال GPT-3 ، GPT-4، Bard، Claude وغيرها من الأدوات الشهيرة.	– إصدار مجاني محدود يتيح استكشافًا على نطاق صغير بدون قدرات تجسيد النصوص. للحصول على وصول أوسع، تبدأ الخطط الشخصية بسعر ٥ دولار شهريًا لتحليل ما يصل إلى ١٠,٠٠٠ كلمة، كما تتوفر استخدام على مستوى المؤسسة من خلال التسعير التجاري المخصص بناءً على الطلب.
GptZero	هي أداة مفتوحة المصدر للكشف عن الذكاء الاصطناعي والتي يمكن استخدامها لاكتشاف المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي في النص. يستخدم GPTZero منهجًا متعدد الخطوات يهدف إلى إنتاج تنبؤات تصل إلى أقصى قدر من الدقة، مع أقل عدد من النتائج الإيجابية الخاطئة. يتخصص النموذج في اكتشاف المحتوى من نماذج Chat GPT و GPT 3 و4 GPT وBard وLLaMa.	– تسمح الطبقة المجانية بمسح ١٠٠٠٠ كلمة شهريًا مع ٧ عمليات فحص في الساعة. – تقوم الخطة الأساسية المدفوعة (١٠ دولارات شهريًا) بمسح ما يصل إلى ١٥٠٠٠٠ كلمة. – يزيد العرض المميز(١٦ دولارًا شهريًا) من الحد الأقصى للكلمات إلى٣٠٠٠٠٠مع إضافة عمليات فحص الانتحال وإمكانيات الدفعات غير المحدودة.

الأداة	الخصائص الرئيسية	نظم التسعير
Crossplag	يجمع CrossPlag بين التعلم الآلي ومعالجة اللغة الطبيعية لتحديد ما إذا كان النص مكتوبًا بواسطة الإنسان أو تم إنشاؤه بواسطة الذكاء الاصطناعي بدقة. ومن خلال تدريبها على مجموعة بيانات واسعة النطاق من المحتوى من مؤلفين بشريين أصليين ونماذج ذكاء اصطناعي، تكشف الأداة عن أنماط لغوية فريدة في كل منهما. من خلال التحليل العميق.	– هناك خطة مجانية محدودة مكونة من ١٠ وحدات دراسية أو ١٠٠٠ كلمة. – تبدأ الخطط المدفوعة من ٩.٩٥ دولارًا أمريكيًا مقابل ٥٠ وحدة دراسية أو ٥٠٠٠ كلمة.
Writer AI	أداة للكشف عن الذكاء الاصطناعي يمكن استخدامها لاكتشاف المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي في النص. يقتصر التحليل على المقتطفات التي يقل طولها عن ١٥٠٠ حرف، لذا تتطلب النصوص الكبيرة عمليات فحص متتالية متعددة. تركز هذه الأداة المبسطة بشكل خاص على قياس مدى اصطناع النص بكفاءة.	– مجاني بالكامل
Sapling	هو أداة كشف الذكاء الاصطناعي التي يمكن استخدامها لاكتشاف المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي في النصوص. تم تصميمه ليتم استخدامه من قبل الأفراد لفحص كتاباتهم عن المحتوى الذي تم إنشاؤه بواسطة الذكاء الاصطناعي. والذي يعد خيارًا جيدًا للأشخاص الذين يبحثون عن أداة مجانية.	– مجاني بالكامل
Winston AI	أداة مساعدة لأي شخص يبحث عن ضمان أصالة المحتوى. تتضمن على خصائص كخيار المسح السريع، وإمكانيات تحميل الملفات، وتقنية التعرف الضوئي على الحروف، وإدارة المشاريع والوثائق، وإدارة الفريق، والدعم متعدد اللغات، وتقارير PDF القابلة للطباعة، وجهاز فحص الاقتباسات.	– تقدم Winston AI مجموعة من خطط التسعير لتناسب احتياجات المستخدمين المختلفة. وتعد الخطة المجانية مثالية لأولئك الذين يحتاجون إلى ميزات أساسية. وتتوفر خطط مدفوعة بالنسبة للميزات المتقدمة. بتكلفة ١٢ دولار شهريًا في الخطة السنوية، ١٨ دولارًا شهريًا للخطة الشهرية. تبلغ تكلفة الخطة النخبة، التي تسمح بفحص ما يصل إلى ٥٠٠،٠٠٠ كلمة، ٤٩ دولارًا شهريًا.

بمراجعة ما ورد من بيانات في الجدول السابق يتبين أن الغالبية العظمي بنسبة ٦٢.٥% من هذه الأدوات مفتوحة المصدر للكشف عن النصوص المصاغة بواسطة أدوات الذكاء الاصطناعي، كما تبين أيضا اعتماد أغلب هذه الأدوات بنسبة ٦٢.٥% على النماذج الشهيرة لفحص النصوص مثل: LLaMa, Bard, GPT 4, GPT 3 ,Claude ، أما فيما يتعلق بنظم تسعيرها، فقد ظهر أن نسبة ٢٥% من الأدوات عينة الدراسة مجانية بالكامل، فيما تبين تقديم ٧٥% منها عدة خيارات للمستخدمين، كالخيارات المجانية، والخيارات المدفوعة، والخيارات المتقدمة، والتي تراوحت بين ١٠ دولارات حتى ١٨ دولار شهريًا.

٣/٣- الجوانب الإيجابية والسلبية لأدوات الدراسة:

ينقل الجدول (٤) الجوانب الإيجابية والجوانب السلبية لأدوات الكشف عن الذكاء الاصطناعي عينة الدراسة، والتي تم الحصول عليها من منصاتها على الويب.

جدول (٤) الجوانب الإيجابية والسلبية لأدوات الكشف عن الذكاء الاصطناعي عينة الدراسة

الأداة	الجوانب الإيجابية	الجوانب السلبية
Copyleaks	- يحدد الشفرة المصدرية المكتوبة بواسطة الذكاء الاصطناعي من أدوات مثل GitHub Copilot ، و ChatGPT في ٣٠ لغة مختلفة مع تلوين النص حسب الكاتب البشري مقابل الكاتب الآلي. - يرصد حتى المحتوى الذكي المنسوخ أو المعاد صياغته ليسلط الضوء على التأثير الجزئي.	- أبطأ في وقت تحميل النتيجة. - يفتقر إلى الاستثناء الجزئي للنص.
Content at scale	- يتضمن على بعض الميزات الرئيسية في طبقته المدفوعة، التي تفتح الكشف غير المحدود للذكاء الاصطناعي عبر النصوص والصور، مع تحديد حتى المحتوى الذي تم إعادة صياغته بواسطة الآلة.	- جودة غير متسقة في إعادة كتابة النصوص بواسطة الذكاء الاصطناعي. - معدلات دقة غامضة وقيود في الكشف.

الأداة	الجوانب الإيجابية	الجوانب السلبية
	– يحصل المستخدمون على ٢,٥٠٠ رصيد إعادة الكتابة paraphrasing لإخفاء مناطق الذكاء الاصطناعي المحددة على أنها كتبت بواسطة الإنسان، – امكانية الكتابة الجماعية المتاحة حسب الطلب. – تشمل القدرات الإضافية مسح محتوى الموقع عبر عناوين.URL . – محرر متميز لإعادة العمل بعد التحليل وتقارير قابلة للتخصيص يتم تصديرها كملفات Word/PDFs/HTML.	
Undetectable AI	– كشف مزدوج بواسطة الذكاء الاصطناعي والإنسان. – يحلل حتى ١٠،٠٠٠ حرف. – تجربة مستخدم بديهية. – يتعرف على أحدث نماذج الذكاء الاصطناعي.	
GptZero	– يوفر GPTZero لوحة معلومات مع عرض تحليلي عميق باستخدام البحث للحصول على لمحات عامة مرئية وإحصائيات دقيقة تكشف عن احتمالية الذكاء الاصطناعي طبقًا للجملة. – يعمل الترميز اللوني بالإضافة إلى التحليل اللغوي القابل للقياس الكمي على تمكين إمكانية الشرح عالي المستوى وإمكانات التحقيق الفني الدقيقة للباحثين والمستخدمين. – يقدم إحصائيات لغوية فريدة من نوعها – يسلط الضوء بصريًا على مناطق النص. – لا يوجد اشتراك، فقط قم بلصق النص.	– مستويات دقة منخفضه نسبيًا. – يفتقر إلى النسبة المئوية للدرجات.

الأداة	الجوانب الإيجابية	الجوانب السلبية
Crossplag	– يتميز CrossPlag بدقة الكشف، وذلك بفضل تشغيله بواسطة نموذج الذكاء الاصطناعي الذي يحتوي على أكثر من ١.٥ مليار معلمة. – يسمح هذا النموذج لـ CrossPlag بالتقاط حتى حالات الانتحال المقنعة بمهارة. – يتيح الذكاء الاصطناعي المتقدم أيضًا أداءً رائدًا في الصناعة في تحديد الفرق بين النص المكتوب بواسطة الإنسان والمحتوى الذي ينشئه الذكاء الاصطناعي. – سهولة الاستخدام.	– التسجيل مطلوب. – لا يسلط الضوء على النص. – معدل دقة منخفض نسبيًا. – محدودية في التعامل مع النصوص الكبيرة/ المعقدة
Writer AI	– يتميز Writer بالمساعدة في الكتابة باستخدام الذكاء الاصطناعي. – التدقيق النحوي/الإملائي. – اكتشاف الانتحال. – توليد الاقتباسات الآلية. – القدرة على تحليل URL المباشر – دعم لغات متعددة - وجميعها تهدف إلى عملية كتابة فعالة وأخلاقية.	– يفتقر إلى الترميز اللوني للذكاء الاصطناعي مقابل النص البشري. – لا يوجد تحميلات للملفات أو خيارات تحليل الصور.
Sapling	– قدرات متقدمة في تحديد المحتوى الذي تم إنشاؤه باستخدام الذكاء الاصطناعي عبر حالات الاستخدام الشائعة والمنصات. – يعتمد خوارزميات التعلم العميق المتقدمة على اكتشاف النص المنشأ بواسطة مساعدي الكتابة الذكية وفحص القواعد النحوية وروبوتات دردشة خدمة العملاء وأدوات المراجعة وموسعات النص وغيرها من الأنظمة التي تستفيد من الذكاء الاصطناعي في إنتاج أو تحسين مسودات الوثائق – سهل الاستخدام نسبياً ومجاني ولا يتطلب التسجيل	– قد لا يكون دقيقًا مثل بعض الأدوات الأخرى.

الأداة	الجوانب الإيجابية	الجوانب السلبية
Winston AI	– دقة عالية. – اكتشاف معظم محتوى الذكاء الاصطناعي، بما في ذلك جميع نصوص GPT-4. – يسلط الضوء على النصوص المولّدة بواسطة الذكاء الاصطناعي. – قادر على معالجة الوثائق المكتوبة يدويًا. – واجهة كشف الذكاء الاصطناعي سهلة الاستخدام.	– يمكن أن يكون الخدمة مكلفة بالنسبة لبعض المستخدمين. – الطبقة المجانية تسمح بما يصل إلى ٢٠٠٠ كلمة شهريًا، مع جلسة دقيقة لا تقل عن ٥٠٠ كلمة. وهذا يعني أنه إذا كنت قد استخدمت بالفعل ١٦٠٠ من رصيد الكلمات الخاص بك، فلن تتمكن من استخدام الـ٤٠٠ المتبقية في الطبقة المجانية. – يتطلب التسجيل للاستخدام.

بتحليل بيانات الجدول السابق يتبين تعدد الميزات التي تقدمها الأدوات عينة الدراسة، والتي أخذت كل أداة منها بإدراجها عبر منصاتها، وإن كانت قد تفاوتت فيما بينها، إلا أن جميعها يركز على سهولة الاستخدام وميزات اكتشاف النصوص، ودعم اللغات المتعددة، وكشف حالات الانتحال المقنع..الخ، كما انحصرت السلبيات في: بطء وقت تحميل النتيجة، الغموض في معدلات الدقة والقيود في الكشف، ضرورة التسجيل للاستخدام، المحدودية في التعامل مع النصوص الكبيرة أو المعقدة..الخ.

٤/٣- دقة التعرف على المقالات المكتوبة بواسطة أدوات الذكاء الاصطناعي:

للتعرف على مستوى دقة أدوات التعرف على الذكاء الاصطناعي عينة الدراسة على النصوص المصاغة بواسطة أدوات الذكاء الاصطناعي، قام الباحث بالاعتماد على أداتين رئيستين لإعداد المقالات العلمية بواسطة الذكاء الاصطناعي، وهما: Writesonic, Google Bard، حيث طلب من Writesonic إعداد مقال علمي حول Freelance Librarians، كما طلب من Google Bard إعداد مقال حول The Internet of Things and Libraries، وعليه تم إخضاع هذين المقالين للفحص بالأدوات عينة الدراسة للوقوف على مستوى الدقة في التعرف على المقالات المكتوبة بواسطة الذكاء الاصطناعي، وينقل الجدول (٥) تفاصيل الفحص.

جدول (٥) دقة التعرف على المقالات المكتوبة بأدوات الذكاء الاصطناعي

الأداة	Google Bard النص الأول					Writesonic النص الثاني				
	اكتشاف النص	تطبيق النسبة	تطبيق الجمل	تصنيف لوني	إدراج تعليقات	اكتشاف النص	تطبيق النسبة	تطبيق الجمل	تصنيف لوني	إدراج تعليقات
Copyleaks	١	٠	٠	١	١	١	٠	٠	١	١
Content at scale	٠	٠	٠	١	٠	٠	٠	٠	١	٠
Undetectable AI	٠	٠	٠	١	٠	١	١	٠	١	١
GptZero	١	١	١	١	١	١	١	١	١	١
Crossplag	١	١	٠	١	١	١	١	٠	١	١
Writer AI	٠	١	٠	٠	٠	٠	١	٠	٠	٠
Sapling	١	١	٠	١	١	١	١	٠	١	١
Winston AI	١	١	١	١	١	١	١	١	١	١
الإجمالي	٦٢.٥ %	٦٢.٥ %	٢٥ %	٨٧.٥ %	٦٢.٥ %	٧٥ %	٧٥ %	٢٥ %	٨٧.٥ %	٧٥ %

بمراجعة الجدول السابق يتبين الارتفاع النسبي لمعدلات التعرف على المقالات المكتوبة بواسطة الذكاء الاصطناعي باستخدام الأدوات عينة الدراسة إجمالاً، كما تبين التفاوت النسبي بين تصنيف المقال الأول والمقال الثاني، حيث تحقق اكتشاف النص الأول بواقع ٦٢.٥% من جملة الأدوات عينة الدراسة في مقابل ٧٥% للنص الثاني، كما جاء تحديد النسبة بواقع ٦٢.٥% في النص الأول في مقابل ٧٥% للنص الثاني من جملة الأدوات عينة الدراسة، أما فيما يتعلق بتحديد الجمل المصاغة بواسطة الذكاء الاصطناعي، فقد صنفتها ٢٥% فقط من جملة أدوات عينة الدراسة في كلا النصين الأول والثاني، كما جاء التصنيف اللوني للنصوص المصاغة بالذكاء الاصطناعي بواقع ٨٧.٥% من جملة الأدوات عينة الدراسة لكلا النصين الأول والثاني، كما تبين تحقيق هذه الأدوات لنسبة ٦٢.٥% في خاصية إدراج التعليقات حول الفحص في النص الأولن في مقابل ٧٥% في النص الثاني.

كما تبين أيضًا التفاوت النسبي بين أدوات الفحص عينة الدراسة في مستوى الدقة في التعرف على النصوص المصاغة بواسطة أدوات الذكاء الاصطناعي، حيث جاءت أداتا: GptZero, Winston AI أفضلها على الإطلاق، في مستوى التعرف على النصوص المصاغة بالذكاء الاصطناعي، وذلك بواقع ١٠٠% من جملة بنود التحليل، بينما جاءت منصتا: content at scale, Writer AI كأدناها على الإطلاق في الإطار ذاته.

وربما تدلنا تلك النتائج على الاختلاف الواضح في مستوى عمق الفحص وعمق تفاصيله بين الأدوات عينة الدراسة، وربما يرجع ذلك بالأساس إلى طبيعة إمكانات هذه الأدوات، ومدى التطوير المستمر في إمكاناتها، ما يجعلها تتفاوت فيما بينها في نتائجها ومؤشراتها.

وتتفق تلك النتائج مع ما توصلت إليه دراسة (Chaka, 2023) من أن أدنى وأعلى نسب احتمالية للنصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي هي ٩٢.٣% و٩٩.٩%، مع وجود ثلاثة نصوص ذات تعادل ٩٩.٩%، كما تم ترتيب أدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة في اكتشافها لنصوص الذكاء الاصطناعي كالتالي: Copyleaks AI Content Detector (1); OpenAI AI Text Classifier (2); GPTZero (3); and Writer.com's AI Content Detector (4), and GLTR (5)

٥/٣ – دقة التعرف على المقالات المكتوبة بواسطة البشر:

للتعرف على مستوى دقة أدوات التعرف على الذكاء الاصطناعي عينة الدراسة في التعرف على المقالات المكتوبة بواسطة البشر، قام الباحث باختيار مقالين علميين مكتوبين من قبل متخصصين في المجال وهما:

1- Lippincott, J. K. (2010). A mobile future for academic libraries. Reference services review, 38(2), 205-213.

2- Ameen, K. (2021), COVID-19 pandemic and role of libraries. Library Management. 42 (4/5), pp. 302-304.

وعليه تم إخضاع هذين المقالين للفحص من قبل أدوات عينة الدراسة لرصد مستوى الدقة في التعرف على المقالات المكتوبة بواسطة البشر، والتفرقة بينها وبين نصوص الذكاء الاصطناعي، وينقل الجدول (٦) تفاصيل الفحص.

جدول (٦) دقة التعرف على المقالات المكتوبة بواسطة البشر

الأداة	المقال الأول					المقال الثاني				
	اكتشاف النص	تحديد النسبة	تحديد الجمل	تصنيف لغوي	إدراج تعليقات	اكتشاف النص	تحديد النسبة	تحديد الجمل	تصنيف لغوي	إدراج تعليقات
Copyleaks	١	٠	٠	٠	١	١	٠	٠	٠	١
Content at scale	١	٠	٠	١	١	١	٠	٠	١	١
Undetectable AI	٠	٠	٠	١	١	١	٠	٠	١	١
GptZero	٠	١	١	١	١	٠	١	١	١	١
Crossplag	١	١	٠	١	١	١	١	٠	١	١
Writer AI	١	١	٠	٠	٠	١	١	٠	٠	٠
Sapling	٠	١	١	١	١	٠	١	١	١	١
Winston AI	١	١	١	١	٠	١	١	١	١	١
الإجمالي	٦٢.٥ %	٦٢.٥ %	٣٧.٥ %	٧٥ %	٧٥ %	٧٥ %	٦٢.٥ %	٣٧.٥ %	٧٥ %	٨٧.٥ %

بتدقيق النظر فيما يحويه الجدول (٦)، يتبين الارتفاع النسبي أيضًا لمعدلات التعرف على المقالات المكتوبة بواسطة البشر باستخدام الأدوات عينة الدراسة، كما تبين التفاوت النسبي بين تصنيف المقال الأول والمقال الثاني، حيث تحقق اكتشاف النص الأول بواقع ٦٢.٥% من جملة الأدوات عينة الدراسة في مقابل ٧٥% للنص الثاني، كما جاء تحديد النسبة بواقع ٦٢.٥% لكلا النصين الأول والثاني من جملة الأدوات عينة الدراسة، ومن الجدير بالذكر، تحديد أداتي: GptZero, Sapling لنسب مصاغة بالذكاء الاصطناعي في كلا النصين الأول والثاني، حيث أشارت

GptZero إلى وجود ٢٥% من النص مصاغ بواسطة الذكاء الاصطناعي، في مقابل ٢% للنص الثاني، كما أشارت أداة Sapling إلى وجود ٠.٢% من النص الأول مصاغ بالذكاء الاصطناعي، في مقابل ٠.١% في النص الثاني، وفيما يتعلق بتحديد الجمل المصاغة بواسطة الذكاء الاصطناعي، فقد صنفتها ٣٧% فقط من جملة الأدوات عينة الدراسة في كلا النصين الأول والثاني، كما جاء التصنيف اللوني للنصوص بواقع ٧٥% من جملة الأدوات عينة الدراسة لكلا النصين الأول والثاني، كما تبين تحقيق هذه الأدوات لنسبة ٧٥% في خاصية إدراج التعليقات حول الفحص في النص الأول في مقابل٨٧.٥% في النص الثاني.

كما تبين أيضًا التفاوت النسبي بين أدوات الفحص عينة الدراسة في مستوى الدقة في التعرف على النصوص المصاغة بواسطة البشر، حيث جاءت أداة: Winston AI أفضلها على الإطلاق، في مستوى التعرف على النصوص المصاغة بواسطة البشر، وذلك بواقع ٩٠% من جملة بنود التحليل، تلتلها في ذلك منصات: GptZero,Crossplag, Sapling، وذلك بواقع ٨٠% من جملة بنود التحليل لكل منها، بينما جاءت منصتا: Copyleaks, Writer AI كأدناها على الإطلاق في ذات الإطار.

وربما تدلنا تلك النتائج على التفاوت بين الأدوات عينة الدراسة في مستوى عمق الفحص للنصوص المصاغة بواسطة البشر، حيث قدمت بعض الأدوات مؤشرات لنصوص مصاغة بواسطة الذكاء الاصطناعي في مقالات حديثة نسبيًا، مما ثير بعض التساؤلات بشأن خوارزميات البحث والمطابقة بهذه الأدوات، وكذا عمق قواعد بياناتها.

ويتفق ذلك مع ما ذهبت إليه دراسة (Dalalah & Dalalah, 2023) من أنه لا توجد حاليًا تقنية قوية تم تطويرها وتقييمها والتحقق منها لتحديد الأكاديميين الذين يستخدمون مولدات نصوص الذكاء الاصطناعي بطريقة غير مناسبة، ولكنها لا تزال بحاجة إلى التحسين؛ فهي عادة ما تقع في فخ الإيجابيات الكاذبة مما يمكن أن يضر مؤلفي العمل الحقيقي، لا سيما في الأدبيات والأجزاء التمهيدية من البحث.

٦/٣- دقة التعرف على مقالات الذكاء الاصطناعي الخاضعة للترجمة:

للتعرف على مستوى دقة أدوات التعرف على الذكاء الاصطناعي عينة الدراسة في رصد المقالات المنشأة بواسطة أدوات الذكاء الاصطناعي، والتي خضعت للترجمة لتغيير هويتها، قام الباحث بإخضاع المقالين اللذين تم صياغتهما سلفًا من قبل Writesonic, Google Bard، كما بالجدول (٥) للترجمة، بالاعتماد على أداتي: sider.ai, openl.io ، حيث تم ترجمة المقال الأول إلى اللغة الألمانية باستخدام sider.ai، والمقال الثاني للغة الفرنسية باستخدام openl.io، ومن ثم تم إخضاع النص المترجم إلى أدوات التعرف على الذكاء الاصطناعي، لتحديد مستوى الدقة في التعرف عليها، وينقل الجدول (٧) تفاصيل الفحص.

جدول (٧) دقة التعرف على مقالات الذكاء الاصطناعي الخاضعة للترجمة

الأداة	الترجمة الألمانية للمقال الأول					الترجمة الفرنسية للمقال الثاني				
	كتبه إنسان	كتبه ذكاء اصطناعي	أعاد صياغته ذكاء اصطناعي	أعاد صياغته إنسان	تم تحريره	كتبه إنسان	كتبه ذكاء اصطناعي	أعاد صياغته ذكاء اصطناعي	أعاد صياغته إنسان	تم تحريره
Copyleaks	٠	٠	٠	٠	٠	٠	٠	٠	٠	٠
Content at scale	٠	٠	٠	٠	١	٠	٠	٠	١	١
Undetectable AI	٠	٠	٠	١	١	٠	٠	٠	١	١
GptZero	٠	١	١	١	١	٠	١	١	١	١
Crossplag	١	١	٠	١	١	٠	١	٠	١	١
Writer AI	-	-	-	-	-	٠	١	٠	٠	١
Sapling	٠	١	١	١	١	٠	١	١	١	١
Winston AI	٠	٠	٠	١	١	٠	٠	٠	١	١
الإجمالي	**١٢.٥ %**	**٣٧.٥ %**	**٢٥ %**	**٦٢.٥ %**	**٧٥ %**	**٠ %**	**٥٠ %**	**٢٥ %**	**٧٥ %**	**١٠٠ %**

يظهر الجدول (٧) الانخفاض النسبي في مستوى التعرف على مقالات الذكاء الاصطناعي الخاضعة للترجمة باستخدام الأدوات عينة الدراسة، كما تبين أيضًا التفاوت النسبي بين الأداوت وبعضها البعض في تصنيف الترجمة الأولى والترجمة

الثانية، حيث تحقق اكتشاف الترجمة الأولى بواقع ١٢.٥% من جملة الأدوات عينة الدراسة في مقابل ٠% للترجمة الثانية، كما جاء تحديد النسبة بواقع ٣٧.٥% للترجمة الأولى، في مقابل ٥٠% للترجمة الثانية، وفيما يتعلق بتحديد الجمل المصاغة بواسطة الذكاء الاصطناعي، فقد صنفتها ٢٥% فقط من جملة الأدوات عينة الدراسة في كلا الترجمتين الأول والثاني، كما جاء التصنيف اللوني للنصوص بواقع ٦٢.٥% للترجمة الأولى في مقابل ٧٥% للترجمة الثانية، كما تبين تميز هذه الأدوات في إدراج التعليقات وذلك بواقع ٧٥% للترجمة الأولى، في مقابل ١٠٠% للترجمة الثانية.

كما تبين أيضًا التفاوت النسبي بين أدوات الفحص عينة الدراسة في مستوى الدقة في التعرف على مقالات الذكاء الاصطناعي الخاضعة للترجمة، حيث جاءت أدوات: Sapling, GptZero كأفضلها على الإطلاق، وذلك بواقع ٨٠% من جملة المعايير لكل منها، ومن الجدير بالذكر أن هذه النسب لم تكن بالضرورة معبرة عن حقيقة النص، وتعكس عمق التعرف عليه، ، حيث أشارت GptZero إلى نسبة ٤٩% ذكاء اصطناعي في الترجمة الأولى، كما أشارت Sapling إلى وجود نسبة ٠.٢% في الترجمة الثانية، في حين أن النص بالكامل مصاغ بالذكاء الاصطناعي، وزيد عليه الترجمة، كما تبين أن Copyleaks كانت أدناها على الإطلاق، والتي لم تجري أي فحص يذكر، لكنها اكتفت بإدراج تعليق "تحتاج لنموذج محسن مدفوع".

وربما تدلنا تلك النتائج على انخفاض مستويات التعرف على النصوص المترجمة من قبل الأدوات عينة الدراسة، بالإضافة إلى انخفاض مستوى عمق الفحص للنصوص المترجمة إجمالاً، أي أن النصوص المترجمة قد لا يمكن اكتشافها بواسطة هذه الأدوات. وتتفق هذه النتائج مع ما ذهب إليه دراسة (Weber-Wulff, et.al, 2023) حيث جاءت الدقة الإجمالية للحالات المكتوبة بشريًا ٩٦%، بينما في حالة الوثائق التي كتبها المؤلفين البشريين بلغات أخرى غير الإنجليزية، ومن ثم تمت ترجمتها آليًا إلى الإنجليزية، انخفضت دقة الأدوات في اكتشافها بنسبة ٢٠%، مما

يشير إلى أن الترجمة الآلية تترك بعض آثار الذكاء الاصطناعي في المخرجات، حتى لو كان النص الأصلي مكتوبًا بواسطة مؤلفين بشريين.

٧/٣ – دقة التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة:

للتعرف على مستوى دقة أدوات التعرف على الذكاء الاصطناعي عينة الدراسة في التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة paraphrasing، قام الباحث بإخضاع المقالين (الأول والثاني) باللغة الانجليزية، لأداتي إعادة الصياغة: quillbot, paraphraser على الترتيب، واللتان تعملان بتقنية الذكاء الاصطناعي، وعليه تم إخضاع المقالين بعد إعادة الصياغة باللغة الإنجليزية للفحص بالأدوات عينة الدراسة لرصد مستوى الدقة في التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة، وينقل الجدول (٨) تفاصيل الفحص.

جدول (٨) دقة التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة

الأداة	المقال الأول بعد إعادة الصياغة بـ quillbot					المقال الأول بعد إعادة الصياغة بـ paraphraser				
	الكشف التام	بخطأ النسبة	بخطأ الجمل	تضليل نحوي	إدراج تعليقات	الكشف التام	بخطأ النسبة	بخطأ الجمل	تضليل نحوي	إدراج تعليقات
Copyleaks	١	١	١	١	١					
Content at scale	٠	٠	٠	٠	١	٠	٠	٠	٠	١
Undetectable AI	٠	١	٠	١	٠	٠	١	٠	١	٠
GptZero	١	١	١	١	١	١	١	١	١	١
Crossplag	١	١	٠	١	١	١	١	٠	١	١
Writer AI	٠	٠	٠	٠	١	٠	٠	٠	٠	١
Sapling	١	١	١	١	١	١	١	١	١	١
Winston AI	٠	٠	٠	٠	١	٠	٠	٠	٠	١
الإجمالي	٥٠ %	٦٢.٥ %	٣٧.٥ %	٦٢.٥ %	٧٥ %	٣٧.٥ %	٥٠ %	٢٥ %	٥٠ %	٦٢.٥ %

يعكس الجدول (٨) الانخفاض النسبي في مستوى التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة أيضًا، باستخدام الأدوات عينة الدراسة، كما تبين التفاوت النسبي بين المنصات وبعضها البعض في تصنيف المقال الأول والمقال الثاني بعد إعادة الصياغة، حيث تحقق اكتشاف صياغة المقال الأول بواقع ٥٠% من جملة الأدوات عينة الدراسة في مقابل ٣٧.٥% لصياغة المقال الثاني، كما جاء تحديد النسبة بواقع ٦٢.٥% للصياغة الأولى، في مقابل ٥٠% للصياغة الثانية، وفيما يتعلق بتحديد الجمل المصاغة بواسطة الذكاء الاصطناعي، فقد صنفتها بواقع ٣٧.٥% للصياغة الأولى في مقابل ٢٥% للصياغة الثانية، كما جاء التصنيف اللوني للنصوص بواقع ٦٢.٥% للصياغة الأولى في مقابل ٥٠% للصياغة الثانية، كما تبين تميز هذه الأدوات إلى حد ما في إدراج التعليقات وذلك بواقع ٧٥% للصياغة الأولى، في مقابل ٦٢.٥% للصياغة الثانية.

كما تبين أيضًا التفاوت النسبي بين أدوات الفحص عينة الدراسة في مستوى الدقة في التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة، حيث جاءت أدوات: GptZero, Sapling كأفلضها على الإطلاق، وذلك بواقع ١٠٠% من جملة المعايير لكل منها، حيث استطاعت هذه الأدوات أن تقدم نسبًا قريبة إلى حد ما من الحقيقة، فقد صنفت GptZero نسبة الذكاء الاصطناعي في الصياغة الأولى بواقع ٩٢%، في مقابل ٨٥% للصياغة الثانية، كما صنفت Sapling ذات النسبة في ٨٧.١% للصياغة الأولى، في مقابل ٧٩% للصياغة الثانية، كما تبين إدراج بعض المنصات لتعليقات تعكس انخفاض مستويات التعرف على النصوص المعاد صياغتها، مثل content at scale والتي أدرجت تعليق "Hard to tell".

وربما تدلنا تلك النتائج على انخفاض مستويات التعرف على النصوص المعاد صياغتها من قبل الأدوات عينة الدراسة، مما يعني عدم قدرة هذه الأدوات على اكتشاف النصوص المعاد صياغتها.

وتتفق هذه النتائج مع ما ذهبت إليه دراسة (Weber-Wulff, et.al, 2023) من أن استخدام الذكاء الاصطناعي لتحويل النصوص الناتجة عن الذكاء الاصطناعي إلى نصوص يعتبرها المصنفون مكتوبة بواسطة الإنسان، بلغت دقة الإجمالية ٢٦%، مما يعني أن معظم النصوص التي تم إنشاؤها بواسطة الذكاء الاصطناعي تظل غير مكتشفة عند إعادة صياغتها آليًا.

٤- الخاتمة:

١/٤ - النتائج:

- لا تبدو أدوات الكشف عن محتوى الذكاء الاصطناعي بمواصفاتها الحالية جاهزة بما يكفي للكشف الدقيق والمقنع عن المحتوى الناتج عن أدوات الذكاء الاصطناعي.

- تعاني معظم أدوات الكشف عن محتوى الذكاء الاصطناعي أوجهًا متعددة للقصور، لاسيما في خوارزمياتها ومستوى تغطية قواعدها.

- اعتماد أغلب أدوات الكشف عن محتوى الذكاء الاصطناعي عينة الدراسة على النماذج الشهيرة لفحص النصوص مثل: LLaMa, Bard, GPT 4, GPT 3 ,Claude، وقد اختلفت نظم تسعيرها، حيث تبين أن ٢٥% من الأدوات عينة الدراسة مجانية بالكامل، فيما تبين تقديم ٧٥% منها عدة خيارات للمستخدمين، كالخيارات المجانية، والخيارات المدفوعة، والخيارات المتقدمة، والتي تراوحت بين ١٠ دوار، ١٨ دولار شهريًا.

- جاءت أبرز سلبيات أدوات الكشف عن محتوى الذكاء الاصطناعي مركزة في بطء وقت تحميل النتيجة، الغموض في معدلات الدقة والقيود في الكشف، ضرورة التسجيل للاستخدام، المحدودية في التعامل مع النصوص الكبيرة أو المعقدة..الخ.

- التفاوت النسبي والاختلاف الواضح في مستوى عمق الفحص وعمق تفاصيله بين الأدوات عينة الدراسة وكذا مستوى الدقة في التعرف على النصوص المصاغة بواسطة أدوات الذكاء الاصطناعي، حيث جاءت أداتي: GptZero, Winston AI أفضلها على الإطلاق.

- التفاوت النسبي بين أدوات الفحص عينة الدراسة في مستوى الدقة في التعرف على النصوص المصاغة بواسطة البشر، حيث جاءت أداة: Winston AI أفضلها على الإطلاق، في مستوى التعرف على النصوص المصاغة بواسطة البشر، وذلك بواقع ٩٠% من جملة بنود التحليل.

- الانخفاض النسبي في مستوى التعرف على مقالات الذكاء الاصطناعي الخاضعة للترجمة باستخدام الأدوات عينة الدراسة، حيث جاءت أدوات: Sapling, GptZero كأفضلها على الإطلاق، وذلك بواقع ٨٠% من جملة المعايير لكل منها، كما جاءت النسب غير معبرة عن حقيقة النص، ولا تعكس عمق التعرف عليه.

- الانخفاض النسبي في مستوى التعرف على مقالات الذكاء الاصطناعي الخاضعة لإعادة الصياغة أيضًا، باستخدام الأدوات عينة الدراسة، حيث جاءت أدوات: GptZero, Sapling كأفضلها على الإطلاق، وذلك بواقع ١٠٠% من جملة المعايير لكل منها، كما تبين إدراج بعض المنصات لتعليقات تعكس عدم قدرتها على فحص النصوص المعاد صياغتها مثل "Hard to tell".

٢/٤ – التوصيات:

- ضرورة اهتمام المؤسسات القائمة على أدوات الكشف عن محتوى الذكاء الاصطناعي بتطوير خصائصها، من خلال تحسين خوارزمياتها للبحث والمطابقة، وتوسيع قواعد بياناتها، بما يحقق قدر أكبر من الفاعلية والاعتمادية عليها.

- ضرورة دمج أدوات اكتشاف الانتحال مثل Turnitin أو Ithenticate في أدوات الكشف عن محتوى الذكاء الاصطناعي.

- ضرورة اتخاذ الجامعات والمؤسسات الأكاديمية قرارات تلزم الطلاب والباحثين بضرورة فحص إنتاجهم العلمي من خلال أدوات الكشف عن محتوى الذكاء الاصطناعي.

- ضرورة اعتماد المؤسسات الأكاديمية على الإصدارات المدفوعة من أدوات الكشف عن محتوى الذكاء الاصطناعي، والبعد عن الإصدارات المجانية.

- ضرورة اتجاه الجامعات والمؤسسات الأكاديمية في مصر والعالم العربي بتطوير أداة للكشف عن محتوى الذكاء الاصطناعي داعمة للمحتوى العربي.

المصادر والمراجع:

- Akram, A. (2023). An Empirical Study of AI Generated Text Detection Tools. *arXiv preprint arXiv:2310.01423*.
- Anders, B. A. (2023). Why ChatGPT is such a big deal for education. *C2C Digital Magazine*, *1*(18), 4.
- Caren, C. (2023). AI writing: The challenge and opportunity in front of education now. Turnitin.
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, *6*(2).
- Cingillioglu, I. (2023). Detecting AI-generated essays: the ChatGPT challenge. *The International Journal of Information and Learning Technology*, *40*(3), 259-268.
- Dalalah, D., & Dalalah, O. M. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *The International Journal of Management Education*, *21*(2), 100822.
- Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, *19* (1), 17.
- Golan, R., Reddy, R., Muthigi, A., & Ramasamy, R. (2023). Artificial intelligence in academic writing: a paradigm-shifting technological advance. *Nature Reviews Urology. 20*, 327-328.
- Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., ... & Wu, Y. (2023). How close is chatgpt to human experts? comparison corpus, evaluation, and detection. *arXiv preprint arXiv:2301.07597*.

- Huang, J., & Tan, M. (2023). The role of ChatGPT in scientific communication: writing better scientific review articles. *American Journal of Cancer Research*, *13*(4), 1148.
- King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, *16*(1), 1-2.
- Kirchenbauer, J., Geiping, J., Wen, Y., Katz, J., Miers, I., & Goldstein, T. (2023). A watermark for large language models. *arXiv preprint arXiv:2301.10226*.
- Lu, N., Liu, S., He, R., & Tang, K. (2023). Large Language Models can be Guided to Evade AI-Generated Text Detection. *arXiv preprint arXiv:2305.10847*.
- Mitchell, E., Lee, Y., Khazatsky, A., Manning, C. D., & Finn, C. (2023). Detectgpt: Zero-shot machine-generated text detection using probability curvature. *arXiv preprint arXiv:2301.11305*.
- Mohammadzadeh, Z., Ausloos, M., & Saeidnia, H. R. (2023). ChatGPT: high-tech plagiarism awaits academic publishing green light. Non-fungible token (NFT) can be a way out. *Library Hi Tech News*.
- Nature, E. (2023). Revealing the value of nature. *English Nature, Peterborough*. available at: https://0810ajy7h-1103-y-https-www-nature-com.mplbci.ekb.eg/
- Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can ai-generated text be reliably detected?. *arXiv preprint arXiv:2303.11156*.
- Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E. L., ... & Norouzi, M. (2022). Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, *35*, 36479-36494.

- Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing?. *Critical care*, *27*(1), 1-5.
- Schechner S. (2023). ChatGPT Ban Lifted in Italy After Data-Privacy Concessions. Wall Street J.available at: https://www.wsj.com/articles/chatgpt-ban-lifted-in-italy-after-data-privacy-concessions-d03d53e7. Accessed 1/1/2024
- Thorp, H. H. (2023). ChatGPT is fun, but not an uthor. *Science*, 379(6630), 313-313.
- Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. *Open Information Science*, *7*(1), 20220158.
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, *19*(1), 26.
- Xiang, H., Zhou, J., & Xie, B. (2023). AI tools for debunking online spam reviews? Trust of younger and older adults in AI detection criteria. Behaviour & Information *Technology*, *42*(5), 478-497.