

عنوان التقرير
نظام ذكي للكشف عن الأخبار المزيفة باستخدام الذكاء الاصطناعي

إعداد الطالب/
طه طارق طقيقه

إشراف/
د عبد الرحيم نادر

العام الدراسي:

2026

الإطار النظري والدراسات السابقة

تمهيد

يتناول هذا الفصل الإطار النظري المرتبط بموضوع البحث، من خلال توضيح مفهوم الأخبار المزيفة وأسباب انتشارها، والتعرف على دور الذكاء الاصطناعي والتعلم الآلي ومعالجة اللغة الطبيعية في الكشف عنها. كما يستعرض الفصل عددًا من الدراسات السابقة التي تناولت استخدام تقنيات الذكاء الاصطناعي والتعلم الآلي والتعلم العميق في تصنيف الأخبار والكشف عن الأخبار المزيفة، بهدف الاستفادة من نتائجها في بناء النظام الذكي المقترح في هذا البحث.

أولاً: مفهوم الأخبار المزيفة

تشير الأخبار المزيفة إلى الأخبار أو المعلومات التي يتم تقديمها بصورة تبدو إخبارية، رغم أنها تتضمن معلومات غير صحيحة أو مضللة، وقد يكون الهدف منها التأثير في آراء المستخدمين أو توجيه قراراتهم. وقد أصبحت هذه المشكلة أكثر وضوحًا مع الانتشار الواسع للإنترنت ووسائل التواصل الاجتماعي، حيث أصبح من السهل إنتاج المحتوى ونشره والوصول إلى عدد كبير من المستخدمين خلال فترة قصيرة.

وتكمن صعوبة التعامل مع الأخبار المزيفة في أن بعض المحتويات المضللة قد تكون مكتوبة بأسلوب مشابه للأخبار الحقيقية، مما يجعل اكتشافها يدويًا أمرًا يحتاج إلى وقت وجهد. ولذلك أصبح الكشف الآلي عن الأخبار المزيفة أحد الموضوعات المهمة في مجال معالجة اللغة الطبيعية والذكاء الاصطناعي.

وقد أوضحت الدراسات في مجال معالجة اللغة الطبيعية أن الكشف الآلي عن الأخبار المزيفة يمثل مشكلة معقدة، وأن الباحثين استخدموا مجموعة متنوعة من البيانات والأساليب والنماذج لمعالجة هذه المشكلة (Oshikawa, Qian, & Wang, 2020).

ثانيًا: أسباب انتشار الأخبار المزيفة

يرتبط انتشار الأخبار المزيفة بعدة عوامل، من أهمها سرعة انتشار المعلومات عبر وسائل التواصل الاجتماعي، وسهولة إعادة نشر المحتوى، بالإضافة إلى اعتماد بعض المستخدمين على العناوين والمحتويات المنتشرة دون التحقق من مصادرها.

كما ساهمت زيادة حجم المحتوى المنشور على الإنترنت في جعل عملية التحقق اليدوي من جميع الأخبار أكثر صعوبة. ولذلك ظهرت الحاجة إلى تطوير أنظمة آلية تستطيع تحليل كميات كبيرة من البيانات خلال وقت قصير، والمساعدة في تحديد الأخبار التي يحتمل أن تكون مزيفة.

وتشير الدراسات الحديثة إلى أن انتشار الأخبار المزيفة عبر المنصات الرقمية يمثل تحديًا مستمرًا، وأن استخدام الذكاء الاصطناعي والتعلم العميق أصبح من الاتجاهات المهمة في التعامل مع هذه المشكلة (Hu et al., 2022).

ثالثًا: الذكاء الاصطناعي والتعلم الآلي

يعد الذكاء الاصطناعي من التقنيات الحديثة التي تتيح للحواسيب والأنظمة البرمجية تنفيذ مهام تتطلب تحليل البيانات والتعرف على الأنماط واتخاذ قرارات مبنية على البيانات.

ويعد التعلم الآلي أحد المجالات الرئيسية في الذكاء الاصطناعي، حيث يعتمد على تدريب النماذج باستخدام بيانات سابقة حتى تتمكن من التعرف على الأنماط واستخدامها في إجراء التصنيف أو التنبؤ عند التعامل مع بيانات جديدة.

وفي مجال الكشف عن الأخبار المزيفة، يمكن تدريب نموذج للتعلم الآلي على مجموعة من الأخبار المصنفة مسبقًا إلى أخبار حقيقية وأخبار مزيفة، ثم استخدام النموذج لتصنيف أخبار جديدة بناءً على الخصائص التي تعلمها من بيانات التدريب.

وقد أظهرت الدراسات المقارنة أن نماذج التعلم الآلي والتعلم العميق يمكن استخدامها في تصنيف الأخبار، مع وجود اختلاف في أداء النماذج بحسب طبيعة البيانات وحجمها وخصائصها. كما أظهرت بعض الدراسات أهمية النماذج اللغوية المدربة مسبقًا، مثل BERT ، في مهام الكشف عن الأخبار المزيفة. (Kaliyar et al., 2021).

رابعًا: معالجة اللغة الطبيعية ودورها في الكشف عن الأخبار المزيفة

تعد معالجة اللغة الطبيعية (Natural Language Processing - NLP) من المجالات المهمة في الذكاء الاصطناعي، وتهدف إلى تمكين الحاسوب من التعامل مع اللغة البشرية وتحليل النصوص واستخراج المعلومات والخصائص منها.

وفي النظام المقترح في هذا البحث، يمكن استخدام تقنيات معالجة اللغة الطبيعية لتحليل محتوى الأخبار، مثل العنوان والنص والكلمات المستخدمة، وتحويل النصوص إلى خصائص يمكن للنموذج الذكي استخدامها في عملية التصنيف.

وقد أكدت الدراسات المتخصصة في هذا المجال أهمية معالجة اللغة الطبيعية في الكشف عن الأخبار المزيفة، كما تناولت هذه الدراسات عددًا من طرق تمثيل النصوص والنماذج المستخدمة في التصنيف، مع الإشارة إلى أن طبيعة الأخبار وتنوعها يمثلان تحديًا أمام بناء نماذج دقيقة وقابلة للتعميم. (Oshikawa et al., 2020).

خامسًا: الدراسات السابقة

الدراسة الأولى (Wang, 2017) :

قدم Wang دراسة بعنوان:

“Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection”

ركزت الدراسة على توفير مجموعة بيانات معيارية يمكن استخدامها في أبحاث الكشف الآلي عن الأخبار المزيفة. وتضمنت مجموعة البيانات 12.8 ألف تصريح قصير تم تصنيفها يدويًا، كما استخدمت الدراسة نماذج تعتمد على النص بالإضافة إلى معلومات إضافية مرتبطة بالبيانات.

وأظهرت الدراسة أهمية وجود مجموعات بيانات مصنفة بصورة جيدة لتدريب واختبار نماذج الكشف عن الأخبار المزيفة، كما بينت إمكانية الاستفادة من دمج المعلومات النصية مع معلومات إضافية لتحسين عملية التصنيف. (Wang, 2017).

وتستفيد الدراسة الحالية من هذه الدراسة في التأكيد على أهمية جمع بيانات تدريب واختبار مناسبة تحتوي على أخبار حقيقية وأخبار مزيفة قبل بناء النموذج الذكي.

الدراسة الثانية (Oshikawa, Qian, and Wang (2020) :

تناولت هذه الدراسة موضوع معالجة اللغة الطبيعية في مجال الكشف عن الأخبار المزيفة، وقدمت مراجعة للطرق والبيانات والنماذج المستخدمة في هذا المجال. وأوضحت الدراسة أن الكشف عن الأخبار المزيفة يمثل مشكلة مهمة وصعبة في معالجة اللغة الطبيعية بسبب ضخامة المحتوى المنشور على الإنترنت وتنوع طرق صياغة الأخبار والمعلومات. كما ناقشت عددًا من الأساليب المستخدمة في تحليل النصوص والكشف عن المحتوى المزيف. وتتفق هذه الدراسة مع البحث الحالي في الاعتماد على معالجة اللغة الطبيعية لتحليل المحتوى النصي للأخبار واستخراج الخصائص التي يمكن استخدامها في تصنيف الأخبار (Oshikawa et al., 2020).

الدراسة الثالثة (Kaliyar et al. (2021) :

تناولت هذه الدراسة مقارنة عدد من نماذج التعلم الآلي والتعلم العميق المستخدمة في الكشف عن الأخبار المزيفة، وذلك باستخدام عدة مجموعات بيانات. وأشارت نتائج الدراسة إلى أن النماذج اللغوية المدربة مسبقًا، ومن بينها BERT، حققت أداءً قويًا في مهمة الكشف عن الأخبار المزيفة، وخاصة عند التعامل مع مجموعات بيانات صغيرة نسبيًا. كما ناقشت الدراسة تأثير خصائص البيانات وطول المقالات والموضوعات على أداء النماذج. وتفيد هذه الدراسة البحث الحالي في اختيار النموذج المناسب ومراعاة طبيعة وحجم البيانات المستخدمة في التدريب والاختبار (Kaliyar et al., 2021).

الدراسة الرابعة (Hu et al. (2022) :

قدمت هذه الدراسة مراجعة شاملة لاستخدام تقنيات التعلم العميق في الكشف عن الأخبار المزيفة. وقامت بتصنيف الأساليب المستخدمة وفق طبيعة الإشراف على التعلم والخصائص المستخدمة، مثل محتوى الأخبار والسياق الاجتماعي والمعرفة الخارجية. كما تناولت الدراسة عددًا من مجموعات البيانات المستخدمة في أبحاث الكشف عن الأخبار المزيفة، وناقشت التحديات والاتجاهات المستقبلية في هذا المجال. وتتفق هذه الدراسة مع البحث الحالي في الاعتماد على تقنيات الذكاء الاصطناعي والتعلم الآلي لتحليل الأخبار وتصنيفها، كما تدعم أهمية تقييم النموذج باستخدام مقاييس علمية مناسبة (Hu et al., 2022).

الدراسة الخامسة (Alnabhan and Branco (2024) :

قدمت هذه الدراسة مراجعة منهجية لأبحاث الكشف عن الأخبار المزيفة التي تعتمد على تقنيات التعلم العميق. وركزت على الخوارزميات المستخدمة ومجموعات البيانات وفعالية الحلول المختلفة، كما ناقشت عددًا من التحديات مثل مشكلة عدم توازن البيانات واستخدام التعلم بالنقل. وتبرز أهمية هذه الدراسة في توضيح التطور المستمر في استخدام التعلم العميق للكشف عن الأخبار المزيفة، وضرورة الاهتمام بجودة البيانات وطريقة تدريب النموذج عند تطوير نظام ذكي للكشف عن الأخبار (Alnabhan & Branco, 2024).

الدراسة السادسة (Alghamdi, Lin, and Luo 2026) :

تناولت هذه الدراسة مراجعة منهجية لأساليب التعلم الآلي والتعلم العميق المستخدمة في الكشف عن الأخبار المزيفة في سياقات متعددة اللغات، مع التركيز على اللغات منخفضة الموارد.

وأشارت الدراسة إلى أن معظم الأبحاث السابقة ركزت بصورة كبيرة على اللغة الإنجليزية، بينما تواجه اللغات الأقل توفرًا للبيانات تحديات تتعلق بقلّة مجموعات البيانات والأدوات المناسبة. كما استعرضت الدراسة الأساليب والبيانات ومقاييس التقييم المستخدمة في هذا المجال.

وتكتسب هذه الدراسة أهمية بالنسبة للبحث الحالي لأنها توضح أن طبيعة اللغة وحجم البيانات المتاحة يمكن أن يؤثرًا في أداء نموذج الكشف عن الأخبار المزيفة، وهو أمر ينبغي أخذه في الاعتبار عند بناء النظام المقترح. (Alghamdi, Lin, & Luo, 2026).

سادسًا: التعقيب على الدراسات السابقة

من خلال مراجعة الدراسات السابقة يتضح أن استخدام الذكاء الاصطناعي والتعلم الآلي والتعلم العميق أصبح من الاتجاهات المهمة في مجال الكشف عن الأخبار المزيفة. كما توضح الدراسات أهمية معالجة النصوص واستخدام مجموعات بيانات مصنفة لتدريب النماذج واختبارها.

وتشير الدراسات كذلك إلى أن أداء النظام لا يعتمد على الخوارزمية المستخدمة فقط، بل يتأثر بجودة البيانات وحجمها وتنوعها، وطريقة تمثيل النصوص، وطبيعة الأخبار المستخدمة في التدريب والاختبار.

ويتميز البحث الحالي بتركيزه على تصميم نظام ذكي مقترح للكشف عن الأخبار المزيفة من خلال تحليل محتوى الأخبار باستخدام تقنيات معالجة اللغة الطبيعية والتعلم الآلي، ثم تقييم أداء النظام باستخدام معايير مثل الدقة (Accuracy) والاستدعاء (Recall) ومعامل F1، وهو ما يتوافق مع المنهجية المحددة في البحث الحالي.

خلاصة الفصل

تناول هذا الفصل مفهوم الأخبار المزيفة وأسباب انتشارها، كما تم توضيح دور الذكاء الاصطناعي والتعلم الآلي ومعالجة اللغة الطبيعية في الكشف عنها. وتم كذلك استعراض عدد من الدراسات السابقة التي تناولت مجموعات البيانات والخوارزميات ونماذج التعلم العميق المستخدمة في هذا المجال.

وتساعد نتائج الدراسات السابقة في بناء الأساس النظري للنظام المقترح، وتحديد الجوانب التي ينبغي مراعاتها عند جمع البيانات ومعالجتها وتدريب النموذج وتقييم أدائه.

المراجع الإضافية

- .1 Alghamdi, J., Lin, Y., & Luo, S. (2026). Machine learning and deep learning approaches for fake news detection and related topics in multilingual contexts: A systematic literature review. *Multimedia Tools and Applications*, 85, 353.
- .2 Alnabhan, M. Q., & Branco, P. (2024). Fake news detection using deep learning: A systematic literature review. *IEEE Access*, 12, 114435–114459.
- .3 Hu, L., Wei, S., Zhao, Z., & Wu, B. (2022). Deep learning for fake news detection: A comprehensive survey. *AI Open*, 3, 133–155. <https://doi.org/10.1016/j.aiopen.2022.09.001>
- .4 Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2021). A benchmark study of machine learning models for online fake news detection. *Decision Support Systems*, 152, 113324.
- .5 Oshikawa, R., Qian, J., & Wang, W. Y. (2020). A survey on natural language processing for fake news detection. *Proceedings of the 12th Language Resources and Evaluation Conference*, 6086–6093.
- .6 Wang, W. Y. (2017). “Liar, Liar Pants on Fire”: A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 422–426.