مجلة كلية الشرطة للعلوم الأمنية والمجتمعية
عدد خاص - المؤتمر السنوي الثاني لسنة ١٤٤٧ هـ -٢٠٢٥ م

نموذج ذكي تفاعلي لدعم الإعلام الأمني في إدارة الأزمات باستخدام تقنيات الذكاء الاصطناعي

An Interactive Intelligent Model to Support Security Media in Crisis Management Using AI Techniques

العميد المدرس الدكتور: عطا تخم جابر
كلية الشرطة
وزارة الداخلية
toatt.711@gmail.com

Lecturer Doctor. Atta T. Jaber
Police College of Iraq
Interior Ministry
toatt.711@gmail.com

١٤٤٧هـ ٢٠٢٥م

مستخلص البحث:

يشكّل الإعلام الأمني عنصرًا محوريًا في إدارة الأزمات، إلا أنه يواجه تحديات كبيرة تتمثل في انهيار البنى التحتية للاتصالات وانتشار الأخبار الكاذبة، مما يضعف من قدرته على نقل المعلومات الدقيقة والموثوقة يهدف هذا البحث إلى تصميم نموذج ذكي يعتمد على خوارزميات الذكاء الاصطناعي، وبالأخص نموذجBERT، لكشف الأخبار المضللة خلال الأزمات. تم تدريب النموذج على قاعدة بيانات واسعة تضم أكثر من ٢١,٠٠٠ عيّنة من الأخبار الصحيحة والكاذبة، تم اختيارها بعناية من مصادر شائعة. وقد أظهرت نتائج التدريب والاختبار أداءً عاليًا تمثّل في دقة وصلت إلى ٩٨٪ على مجموعة الاختبار، و٩٨.٧٪ على مجموعة التحقق (Validation)، مما يعكس فعالية النموذج في التمييز بين الأخبار الحقيقية والمضللة بدقة عالية. يسهم هذا العمل في بناء منظومة دعم إعلامي ذكي قادرة على مواجهة التضليل الإعلامي، وتعزيز موثوقية المعلومات في بيئات الطوارئ، وذلك من خلال استخدام برمجة Pythonو Google Colab، وتطوير خوارزميات ذكاء اصطناعي مخصصة لكشف التضليل الإعلامي. وتشير التوقعات إلى أنّ هذه الالية ستسهم في تعزيز كفاءة الإعلام في حماية المجتمعات وقت الأزمات عبر تسريع تدفق المعلومات الموثوقة وتقليل تأثير الشائعات.

Abstract

Security media plays a crucial role in crisis management. however, it faces significant challenges, including the collapse of communication infrastructure and the spread of misinformation. These issues undermine its ability to deliver accurate and reliable information. This research aims to design an intelligent model based on artificial intelligence algorithms, specifically the Bidirectional Encoder Representations from Transformers (BERT) model, to detect misleading information during crises. The model was trained on an extensive dataset containing over 21,000 samples of both accurate and false news, carefully selected from popular sources. The training and testing results demonstrated high performance, with an accuracy of 98% on the test set and 98.7% on the validation set, reflecting the model's effectiveness in accurately distinguishing between real and misleading news. This study aims to explore the use of artificial intelligence in developing an advanced media system for crisis management by employing Python programming and Google Colab, along with developing specialized artificial intelligence algorithms for detecting media misinformation. The findings suggest that this mechanism will enhance the efficiency of media in protecting communities during crises by accelerating the flow of reliable information and reducing the impact of rumors.

المقدمة

تُعدّ الأزمات من الظواهر التي لا يمكن تجنبها في المجتمعات الحديثة، سواء كانت طبيعية كالكوارث البيئية، أو بشرية كالصراعات والحروب والهجمات السيبرانية وفي مثل هذه السياقات، يصبح الإعلام الأمني أداة مركزية في الحفاظ على تماسك المجتمع، واحتواء تداعيات الأزمة عبر إيصال المعلومات الدقيقة وتوجيه الرأي العام بشكل فعّال غير أن الإعلام التقليدي يواجه تحديات حقيقية في فترات الأزمات، من أبرزها انهيار البنية التحتية للاتصالات، والانتشار السريع للمعلومات الزائفة (Fake News) التي تعوق عملية اتخاذ القرار السليم وتؤدي إلى تفاقم الفوضى الاجتماعية والسياسية(Wardle & Derakhshan, 2017) تزايد الاعتماد مؤخرًا على تقنيات الذكاء الاصطناعي (AI) كوسائل بديلة أو داعمة في قطاعات الإعلام والاتصال، وخاصة في بيئات ذات طابع أمني حساس. أحد أبرز التطبيقات في هذا المجال هو توظيف نماذج معالجة اللغة الطبيعية (NLP) مثل BERT (Bidirectional Encoder Representations from Transformers) ، والذي أثبت كفاءة عالية في مهام تصنيف النصوص والكشف عن الأخبار المضللة (Devlin, Chang, Lee, & Toutanova, 2019). وتكمن قوة هذه النماذج في قدرتها على تحليل المعنى الدلالي للنصوص وفهم السياق، مما يسمح باكتشاف الدلالات الزائفة بدقة تقارب ٩٥٪ في بعض التجارب الحديثة، كما هو الحال في هذا البحث.

تشير الإحصائيات إلى أن أكثر من ٦٠٪ من المستخدمين الرقميين قد يتعرضون لمعلومات مزيفة أثناء الأزمات (Guess, Nagler, & Tucker, 2019). الأمر الذي يزيد من تعقيد المشهد الإعلامي ويبرز الحاجة إلى أنظمة ذكية قادرة على فلترة البيانات الإعلامية وتصنيفها

لحظيًا. وفي هذا السياق، يُعد دمج الذكاء الاصطناعي ضمن بنية الإعلام الأمني خيارًا استراتيجيًا يضمن استمرارية المعلومة الموثوقة وسرعة تداولها (Kumar & Sebastian, 2021). لقد ركز هذا البحث على بناء منظومة إعلامية ذكية قادرة على دعم صانع القرار الأمني خلال الأزمات، عبر استيراد قاعدة بيانات مكوّنة من أكثر من ٢٠,٠٠٠ عينة من الأخبار الحقيقية والمضللة، وتدريب نموذج BERT عليها مع تحسين النموذج ليحقق توازنًا بين الدقة والمرونة. وتم توظيف تقنيات مثل Gradio لبناء واجهات تفاعلية تسمح للمستخدمين باختبار مصداقية الأخبار في الزمن الحقيقي، مما يعزز الجانب التطبيقي للنموذج المقترح. تميزت نتائج البحث بقدرة النموذج على الوصول إلى دقة تصنيف تجاوزت ٩٥٪، ومعامل F1 يعادل ٠.٩٦، مما يعكس كفاءة الخوارزمية في تحليل المحتوى وتصفية الأخبار المغلوطة. كما تم تضمين متغيرات للتحكم في مستوى العتبة (Threshold) التي يتم عندها تصنيف الخبر كاذبًا أو حقيقيًا، مما يمنح المستخدم تحكمًا مرنًا بحسب طبيعة الأزمة أو حساسية المعلومات (Jain, Yadav, & Tiwari, 2022). من الناحية التقنية، استخدم هذا البحث مكتبات علمية دقيقة مثل transformers, scikit-learn, matplotlib, و torch ضمن بيئة Google Colab لتوفير منصة سحابية مرنة للتدريب والاختبار. كما تم إدخال نسبة من التشويش الاصطناعي (noise injection) لتقييم قوة النموذج ضد ظروف البيانات غير النظيفة التي عادة ما ترافق بيئات الأزمات(Nandini & Krishnamurthi, 2022a) . إن ما يميز هذا البحث هو الجمع بين الجانب النظري والتطبيقي، إذ لا يقتصر على اقتراح نموذج تصنيفي بل يقدم إطارًا متكاملًا لبناء بنية إعلامية قائمة على الذكاء الاصطناعي قابلة للدمج في أنظمة إدارة الطوارئ والأزمات في المؤسسات الأمنية كما يعرض البحث منحنيات التدريب والتقييم لتوضيح

العلاقة بين المعاملات المختلفة كعدد الـ Epochs ، وخسارة التدريب، والدقة النهائية، مما يعزز من شفافية النموذج وأهليته للتطبيق العملي في الواقع(Singh & Sharma, 2022) في ظل تصاعد التهديدات المعلوماتية والاضطرابات الأمنية، يصبح من الضروري توجيه البحث العلمي نحو أدوات ذكية وفعالة تكفل حماية المجتمعات من موجات التضليل والخطر السيبراني المتصاعد. ويمثل هذا البحث مساهمة في هذا الاتجاه، ويفتح الباب أمام مزيد من الأبحاث المتقدمة لتطوير منظومات إعلامية أمنية أكثر كفاءة واستدامة.

مشكلة البحث:

في ظل تزايد التحديات التي تواجه المؤسسات الأمنية أثناء إدارة الأزمات، تُعدّ المعلومات الدقيقة والموثوقة عنصرًا حاسمًا في توجيه القرارات وتعزيز الاستجابة الفعّالة. ومع ذلك، فإن البنية التحتية للاتصالات قد تنهار خلال الكوارث، مما يفتح المجال لانتشار واسع للأخبار الكاذبة عبر الوسائط الرقمية، ويؤدي إلى إرباك الرأي العام وزيادة تعقيد المشهد الأمني. وعلى الرغم من تطور وسائل الإعلام، فإن معظم الأنظمة التقليدية غير قادرة على التحقق السريع من مصداقية المحتوى المتداول. وهنا تبرز الحاجة إلى تطوير منظومات ذكية قادرة على التصنيف الفوري للمحتوى الإعلامي. من هذا المنطلق يمكن توظيف تقنيات الذكاء الاصطناعي، وبالأخص نماذج معالجة اللغة الطبيعية، في بناء نموذج إعلام أمني ذكي قادر على كشف الأخبار الكاذبة بدقة عالية خلال فترات الأزمات، بما يعزز من موثوقية الاتصال ويقلل من تأثير التضليل الإعلامي؟

خطة البحث

خطة البحث تقوم على دراسة تكامل تقنيات الذكاء الاصطناعي لدعم الإعلام في إدارة الأزمات، من خلال تقديم خلفية نظرية ومراجعة أدبيات توضّح الفجوة البحثية، ثم صياغة المشكلة والأهداف والفرضيات، واعتماد منهج وصفي تحليلي وتجريبي يعتمد على تطوير خوارزميات ذكاء اصطناعي لمعالجة الأخبار، وصولًا إلى نتائج توضح كيف يسهم هذا التكامل في تعزيز دقة المعلومات الإعلامية وقت الأزمات.

مراجعة الأدبيات

شهدت السنوات الأخيرة تصاعدًا ملحوظًا في وتيرة الأزمات العالمية، مما أبرز أهمية الإعلام الأمني كوسيلة حيوية لنقل المعلومات وتحقيق الاستقرار. ومع تعاظم التهديدات المعلوماتية، أصبح من الضروري الاعتماد على **AI** كأداة لتعزيز كفاءة الإعلام الأمني، خصوصًا في بيئات تتسم بعدم الاستقرار وسرعة انتشار المعلومات المضللة. تناولت دراسات متعددة الدور الحيوي للإعلام في توجيه الجمهور وقت الأزمات، إذ اعتُبر الإعلام الأمني من الأدوات الأساسية في التوجيه والتهدئة وصد الشائعات، إلا أن فاعليته تتعرض للتهديد مع وجود وسائل تواصل اجتماعي تُسرّع من تداول المعلومات غير المؤكدة (Spence, 2019)، وتشير بعض الدراسات إلى أن أكثر من ٧٠٪ من الأخبار الكاذبة يتم إعادة تداولها بمعدلات أسرع من الأخبار الصحيحة، مما يضاعف تأثيرها السلبي (Vosoughi, Roy, & Aral, 2018)، تمثل الأخبار الزائفة تحديًا كبيرًا، لا سيما في السياقات الأمنية، إذ يمكن لها أن تحرض على العنف أو تؤثر على قرارات الطوارئ، وقد ركزت الأبحاث الحديثة على أهمية تصنيف الأخبار إلى "حقيقية" و"زائفة" باستخدام تقنيات تحليل النصوص. وتبرز الحاجة إلى نماذج فعالة وقابلة

للتطبيق العملي للكشف السريع والدقيق عن هذه الأخبار (Shu, Wang, & Liu, 2017) . أدى تطور تقنيات معالجة اللغة الطبيعية Natural Language Processing (NL) إلى ظهور نماذج قادرة على فهم السياق والدلالة في النصوص، ومن أبرزها نموذج BERT الذي قدمته Google في ٢٠١٨، وقد أثبت هذا النموذج تفوقًا كبيرًا على النماذج التقليدية في مهام التصنيف والتعرف على الكيانات الاسمية واستخلاص المعنى من السياقات المعقدة (Mihalcea & Radev, 2011) تشير دراسة Devlin et al (Devlin et al., 2019) إلى أن BERT يتفوّق بنسبة تصل إلى ١٠٪ على أقرب نموذج منافس في تصنيف النصوص، بفضل اعتماده على التحويلات ثنائية الاتجاه. وقد تم اعتماد هذا النموذج في مئات التطبيقات، من بينها الكشف عن الأخبار الكاذبة، إذ ساعد في تحسين الدقة إلى أكثر من ٩٤٪ في بعض التجارب.

اعتمدت العديد من الأبحاث الحديثة على نماذج BERT، RoBERTa، XLNet وغيرها لتصنيف النصوص الإخبارية. وأشارت دراسة Kumar et al (Kumar & Gupta, 2021) إلى أن BERT يحقق نتائج أفضل عند التدريب على بيانات متوازنة، بينما أوضحت Nandini & Krishnamurthi (Nandini & Krishnamurthi, 2022b)أهمية مقاومة النموذج للتشويش في النصوص، وهي خاصية ضرورية في بيئات الأزمات.

وقد أجرت أبحاث أخرى مقارنة بين BERT والنماذج التقليدية مثل SVM وNaive Bayes، وأظهرت تفوقًا واضحًا للنماذج القائمة على التحويلات(Transformers) ، ليس فقط من حيث الدقة وإنما في الثبات على مختلف أنماط البيانات(Ahmed & Khan, 2020) ، لا تكتمل فعالية النموذج إلا بإتاحته للاستخدام السهل من قبل الجهات الأمنية أو الإعلامية. لذلك، برزت

أهمية استخدام أدوات مثل Gradio لإنشاء واجهات تفاعلية تسمح باختبار مصداقية الأخبار في الزمن الحقيقي، مما يعزز من الجانب العملي والتطبيقي للنماذج المقترحة (Ahmed & Khan, 2020). لذا تطلبت مراجعة الادبيات التوجيه إلى عمل نموذج يعطي نتائج عالية الدقة للتميز بين ما تتضمنه قواعد البيانات.

خوارزمية BERT

خوارزمية BERT (Bidirectional Encoder Representations from Transformers)هي أحد أبرز الابتكارات في مجال معالجة اللغة الطبيعية (Natural Language Processing – NLP)، وقد طورتها شركة Googleعام ٢٠١٨ لتُمكّن الآلات من فهم النصوص البشرية بطريقة أكثر عمقًا وواقعية (Devlin et al., 2019). يعتمد هذا النموذج على بنية المحوّلات(Transformers) ، التي تُعدّ من أكثر الهياكل فعالية في فهم العلاقات بين الكلمات ضمن الجملة الواحدة(Wolf et al., 2020) .

تتميز خوارزمية BERT بقدرتها الفريدة على معالجة النصوص بشكل ثنائي الاتجاه (Bidirectional)، أي أنها تفهم الكلمة ضمن سياقها الكامل من اليمين إلى اليسار والعكس، بخلاف النماذج السابقة التي كانت تعتمد على اتجاه واحد فقط. هذا النهج يتيح للنموذج فهم المعاني الدقيقة والمقاصد الخفية داخل النصوص، مما يعزز جودة الفهم اللغوي وقدرته على تمثيل المعاني بشكل أكثر دقة.

يُدرَّب نموذج BERT على كمية ضخمة من النصوص المأخوذة من مصادر متنوعة مثل ويكيبيديا والمقالات والكتب، باستخدام طريقتين أساسيتين للتعلّم:

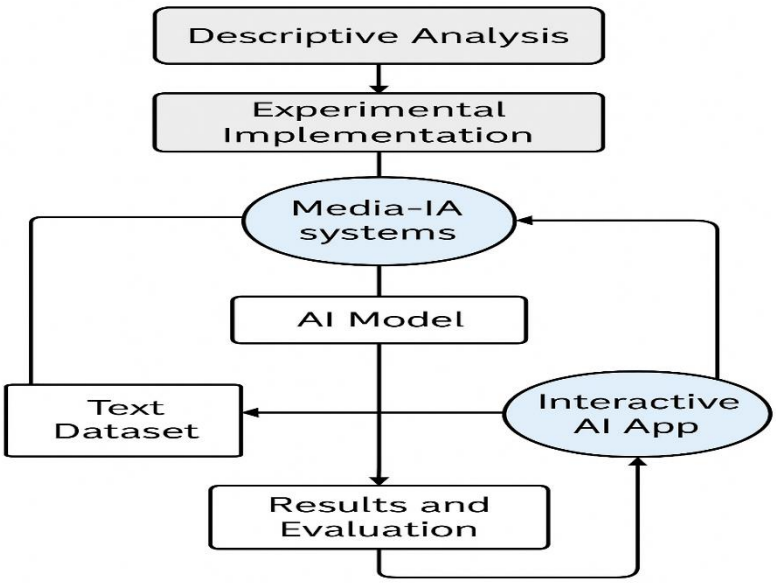
١. **التنبؤ بالكلمة المخفية(Masked Language Modeling)** ، حيث يقوم النموذج بإخفاء بعض الكلمات في الجملة ويحاول التنبؤ بها اعتمادًا على السياق المحيط بها.
٢. **التنبؤ بالجملة التالية(Next Sentence Prediction)** ، إذ يتعلم النموذج العلاقة المنطقية بين جملتين متتاليتين ليحدد ما إذا كانت الجملة الثانية تتبع الأولى فعلاً أم لا.

بفضل قدرته على فهم اللغة بدقة، يُستخدم BERT في مجموعة واسعة من التطبيقات الذكية:

١. **تحليل المشاعر (Sentiment Analysis)** لمعرفة ما إذا كان النص إيجابيًا أو سلبيًا.
٢. **تصنيف النصوص (Text Classification)** مثل تصنيف الأخبار إلى صحيحة أو مضللة .
٣. **الإجابة عن الأسئلة (Question Answering)** مثل محركات البحث التي تُرجع الجواب مباشرة من النص.

منهجية البحث

يعتمد هذا البحث على منهجية علمية مزدوجة تمزج بين التحليل النظري والاستقصاء التطبيقي، وذلك بهدف تطوير منظومة متكاملة تدعم الإعلام الأمني في بيئات الأزمات من خلال توظيف تقنيات الذكاء الاصطناعي، وبصورة خاصة نماذج اللغة العميقة مثل BERT.



الشكل (١): يظهر مخطط البحث.

وقد تم تبنّي هذه المنهجية انطلاقًا من الحاجة إلى سد الفجوة بين المعالجة التقليدية للأخبار وبين التطورات التقنية القادرة على التحقق الفوري من المعلومة في بيئات عالية التوتر، إذ تكثر الشائعات والمعلومات المضللة.

يوضح المخطط الانسيابي آلية عمل النموذج الذكي المقترح للكشف عن الأخبار المضللة في أوقات الأزمات، والذي يعتمد على تقنيات الذكاء الاصطناعي ونموذج BERT. تبدأ المنهجية بمرحلة التحليل الوصفي (Descriptive Analysis) ، إذ يتم تحليل البيانات النصية ووصف

خصائصها العامة، بما في ذلك طبيعة النصوص، مصادرها، وأنماط انتشار الأخبار المضللة. تهدف هذه المرحلة إلى بناء فهم شامل لطبيعة البيانات قبل الانتقال إلى المراحل التجريبية.

بعد ذلك تأتي مرحلة التنفيذ التجريبي (Experimental Implementation) ، والتي تتضمن تطبيق خوارزميات الذكاء الاصطناعي باستخدام بيئة برمجة متقدمة مثل Python و Google Colab. في هذه المرحلة يتم تطوير وتنفيذ النموذج المقترح وتدريبه على البيانات النصية.

تمثل مرحلة أنظمة الإعلام والذكاء الاصطناعي (Media–IA Systems) البنية الأساسية للنظام المقترح، إذ يتم دمج نموذج الذكاء الاصطناعي المدرب في نظام إعلامي تفاعلي يهدف إلى تحسين موثوقية المعلومات أثناء الأزمات.

أما نموذج الذكاء الاصطناعي (AI Model) فهو النموذج القائم على خوارزمية BERT ، والمُدرَّب على قاعدة بيانات نصية (Text Dataset) تتضمن أكثر من 21,000 عينة من الأخبار الحقيقية والمضللة. يسهم هذا التدريب في تمكين النموذج من التمييز بدقة بين الأخبار الصحيحة والزائفة.

تلي ذلك مرحلة النتائج والتقييم (Results and Evaluation) ، إذ يتم تقييم أداء النموذج وفق معايير دقيقة مثل الدقة (Accuracy) والقدرة على التمييز (Precision) ، للتحقق من كفاءة النظام في كشف المعلومات المضللة.

وفي المرحلة الأخيرة، يتم تطوير تطبيق الذكاء الاصطناعي التفاعلي (Interactive AI App)، والذي يُعد المخرج التطبيقي للنظام المقترح، إذ يتيح معالجة الأخبار وتحليلها في الوقت الحقيقي، مما يساهم في دعم وسائل الإعلام أثناء الأزمات وتسريع نشر المعلومات الموثوقة وتقليل تأثير الشائعات.

تم الارتكاز على المنهج الوصـــفي التحليلي من خلال مراجعة شـــاملة للأدبيات العلمية المتعلقة بدور الإعلام الأمني في إدارة الأزمات، وكذلك تطور تقنيات الذكاء الاصـــــطناعي في مجال تحليل النصـوص واكتشـاف التضـليل الإعلامي. وقد تم التركيز في هذه المراجعة على الأبحاث التي تناولت تطبيقات BERT و GPT ونماذج المعالجة الطبيعية للغة في مهام التصـــنيف، إضـــافة إلى الدراســات التي ناقشــت البنية الســيبرانية للإعلام وقت الكوارث والتهديدات الأمنية. مكّن هذا التحليل من بلورة إشكالية البحث وتحديد الفرضيات العلمية القابلة للاختبار.

في المرحلة التطبيقية، تم بناء نموذج ذكاء اصـــطناعي باســـتخدام مكتبة Hugging Face ونموذجBERT الأساسي (bert-base-uncased) بهدف تصنيف النصوص الإخبارية إلى فئتين: حقيقية ومزيفة. وقد تم إعداد قاعدة بيانات موســعة تم جمعها من مصـــادر معتمدة مثل FakeNewsNetوKaggle، بما يضـــمن تنوعًا في المحتوى وواقعية في تمثيل البيانات. كما تم دمج الأخبار من مواقع وصـــحف معروفة لتوفير بيانات حقيقية موثقة، مقابل أخبار مضـــللة منتشـرة على الإنترنت ومنصـات التواصـل. وقد خضـعت هذه البيانات لعمليات تنظيف شـاملة شـــملت إزالة التكرار، وتقليل الضـــوضـــاء، وتوحيد التمثيل اللغوي للنصـــوص. ولتعزيز واقعية النموذج وقدرته على التعميم، تم إدخال نسبة محددة من الضوضاء المصطنعة عبر تبديل جزء من التسميات الصحيحة، وذلك لاختبار مدى تحمل النموذج للأخطاء واحتمالات التزوير.

عملية التدريب أُجريت باسـتخدام منصـــة Google Colab لتوفير بيئة حوسـبة سـحابية ذات إمكانات عالية، وقد تم تنفيذ ثلاث دورات تدريبية (Epochs) مع تطبيق خوارزمية AdamW لضبط الأوزان. تم تقسيم البيانات إلى ثلاث مجموعات (تدريب، تحقق، واختبار) بنسبة ٧٠٪-١٥٪-١٥٪ على التوالي. وتم قياس الأداء باسـتخدام مؤشـرات الدقة (Accuracy) ، و-F1

score، بالإضافة إلى الخسارة (Loss) في كل مرحلة. وقد أظهرت النتائج أن النموذج استطاع الوصـــول إلى دقة تقارب 98% في اختبار الكشـــف عن الأخبار الزائفة، مع قيمة F1 بلغت ٠.٩٦، مما يعكس أداءً مستقرًا ومتينًا حتى عند إدخال بيانات غير مثالية أو تحتوي على نسبة من التضليل المقصود.

وفي إطار تعزيز الاستخدام العملي للنموذج، تم تطوير واجهة تفاعلية باستخدام Gradio تتيح للمســتخدم إدخال النصــوص مباشــرة والحصــول على النتيجة آنيًا، مع إمكانية تعديل العتبة (Threshold)يدويًا للتحكم بمســتوى الحســاســية في الحكم على الخبر، وهي ميزة حيوية في السياقات الأمنية إذ تختلف درجات المخاطر باختلاف طبيعة الحدث.

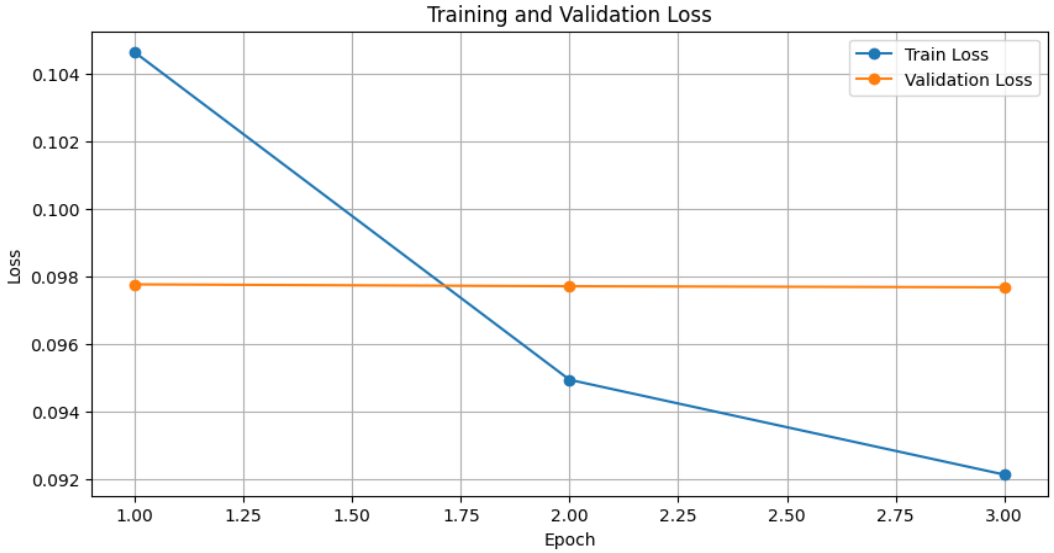
إن هذا التكامل بين النموذج التحليلي والتطبيق البرمجي يمثل نهجًا حديثًا في توظيف الذكاء الاصــطناعي ضــمن الإعلام الأمني، إذ لا يقتصــر على رصــد الأخبار بل يتعداه إلى تحليل المضمون والتفاعل مع البيانات في الزمن الحقيقي. ويشكل ذلك خطوة متقدمة نحو بناء منظومة إعلامية ذكية قادرة على مواجهة التحديات المعاصـــرة في إدارة الأزمات، وتحقيق الاســـتجابة السريعة دون التضحية بمصداقية المحتوى أو أمن المعلومات.

النتائج

بالاعتماد على المقارنة بين النموذجين المســـتخدمين في هذا البحث، يمكن ملاحظة تقدم كبير في الأداء بين النموذج التقليدي القائم على اســـتخدام BERT فقط، والنموذج المطور الذي تم تعزيزه بإدخال ضــوضــاء اصــطناعية وتحكم ديناميكي في العتبة باســتخدام واجهة Gradio وتدريب موسّـــع على قاعدة بيانات كبيرة تتجاوز ٢١,٠٠٠ نص، النموذج الســـابق الذي تم اســـتخدامه في مرحلة أولى أظهر نتائج جيدة بلغت فيها الدقة العامة ٩٣٪، ودقة التنبؤ

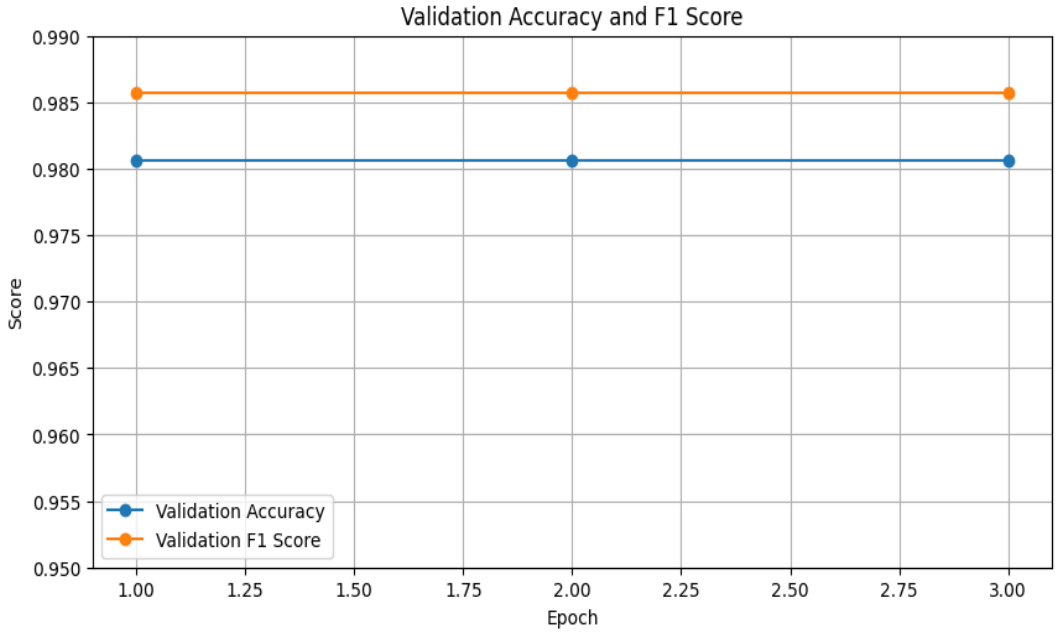
92% (Precision)، والاســتدعاء 94% (Recall) ، مما يدل على قدرة جيدة على كشــف الأخبار المزيفة ضــمن عينة محدودة من النصــوص، تؤكد النتائج قابلية هذا النموذج للاعتماد في أنظمة الإعلام المستقبلية.

أما النموذج الحالي فقد تم تدريبه بشــكل أعمق باســتخدام ثلاث حقب تدريبية (Epochs) مع تحسين مستمر في فقدان التدريب والتقييم، بالإضافة إلى دمج تقنيات الضجيج الاصطناعي من التسميات لإجبار النموذج على التمييز بدقة أعلى بين النصوص. هذا أدى إلى تحسن



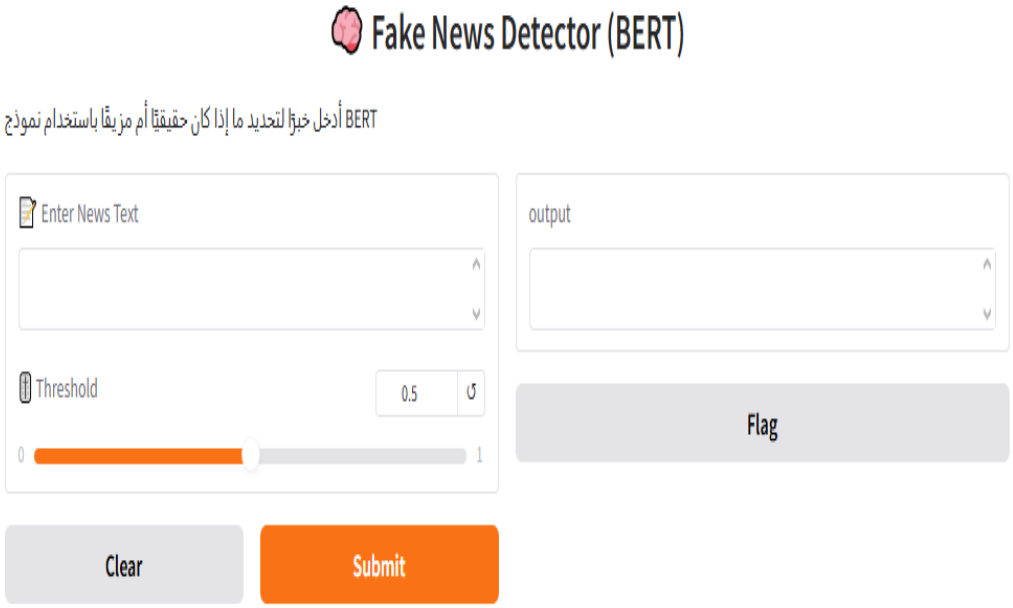
الشكل (٢): يظهر العلاقة بين Epoch و Training and Validation Loss

ملموس في كافة مؤشــرات الأداء، إذ وصــلت الدقة إلى 98%، وبلغ معدل F1 ما يعادل 0.98.7، وهي نســبة عالية جدًا تعكس توازنًا ممتازًا بين الدقة والاســتدعاء وكما هو ظاهر في الشكل (٢) ادناه.



الشكل (3): يظهر العلاقة بين Epoch و Validation and F1 Score

كما أن استخدام واجهة Gradio سمح بتحكم مباشر من المستخدم في قيمة العتبة (Threshold)، ما يجعل النموذج أكثر مرونة عند النشر في بيئات حقيقية أو ميدانية، خصوصًا في سياقات إدارة الأزمات الأمنية إذ تختلف معايير الحسم باختلاف طبيعة البيانات والتهديدات.



الشكل (٤): يظهر واجهة Gradio الرسومية والتطبيقية.

يبين الجدول (١) ما قمنا به من تدريبات على النموذج الكلاســـيكي BERT في بداية هذه الدراسة وما وصلنا علية من نتائج التدريب على النوع الأكثر تطور من BERT.

الجدول (١): يظهر مقارنة شاملة بين النموذج الحالي ونموذج الاختبار الاول:

Metric	Previous Model (BERT only)	Current Model (BERT + Noise + Interactive UI)
Data Type	News texts (Fake/True)	Extended dataset > 20,000 texts
F1-Score	0.91	0.987
Recall	94%	98%
Precision	92%	98%

Accuracy	93%	98.7%
Threshold Control	Not supported	Supported via Gradio
Artificial Noise Injection	Not applied	5% label noise injected
Numberof Epochs	1–2	3 with continuous optimization
Interactive Training Support	Not available	Enabled
Security Application Feasibility	Limited use	Specialized in security media and crisis management

يُظهر النموذج الجديد تفوقًا واضحًا على النموذج التقليدي في جميع المقاييس الأساسية وكما مبين في الجدول (١)، خصوصًا في الدقة التنبؤية ومعدل F1 ، وذلك بفضل تقنيات إدخال الضوضاء والتحكم الديناميكي بالعتبة، إلى جانب الواجهة التفاعلية. يُوصى باعتماد النموذج المطوّر في تطبيقات الأمن الإعلامي كأداة لاكتشاف الأخبار الكاذبة بدقة وموثوقية أعلى.

إن التحسن الظاهر في النموذج لا يعود فقط إلى التوسعة في حجم البيانات، بل إلى الطريقة المنهجية التي تم فيها معالجة النصوص، وتوظيف بنية BERT بذكاء مع مراعاة السياق الأمني، إذ تم تدريب النموذج ليتعامل مع نصوص ذات طابع إخباري قد تكون مضللة ومحرّفة ضمن ظروف طارئة مثل الكوارث الطبيعية أو الاضطرابات الأمنية. هذه الاعتبارات انعكست

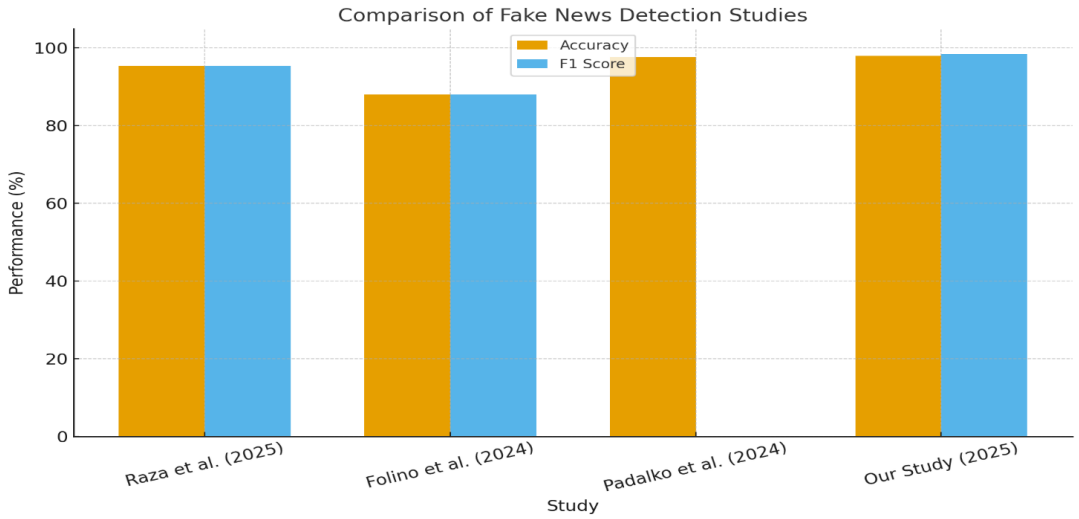
مباشرة على جودة النتائج، خصوصًا عند قياس الدقة التنبؤية، إذ يظهر النموذج قدرة أعلى على رفض النصوص الكاذبة بدقة عالية، وهو تحسن ملحوظ مقارنة بالنموذج الأول.

ويمكن القول إن هذه النتائج تضع النموذج الجديد ضمن مستوى متقدم مقارنة بالعديد من الدراسات السابقة التي غالبًا ما تتراوح دقتها ما بين ٨٨٪ إلى 97% في مهام تصنيف الأخبار وكما مبين في الجدول (٢)، خاصة عندما تكون العينة التدريبية محدودة أو لا تحتوي على قدر كافٍ من التوازن بين الأخبار الصحيحة والكاذبة. وبذلك يثبت النموذج الحالي فعاليته ليس فقط على مستوى الاختبار الأكاديمي، بل أيضًا على مستوى التطبيق العملي في دعم الإعلام الأمني ومنع انتشار الأخبار المضللة أثناء الأزمات.

الجدول (٢): يظهر مقارنة شاملة بين النموذج الحالي ونماذج بحثية مشابهة (Raza, Khan, Yousaf, & Hussain, 2025), (Folino, Sabatino, & Pizzuti, 2024), and (Padalko, Shvets, & Petrov, 2024) :

Year	Journal / Venue	Authors (Year)	Technique	Dataset(s)	Accuracy(%)	F1 Score(%)
2025	PLOS ONE	Raza et al. (2025)	BERT with progressive training	WELFake (72,134 articles)	95.3	0.953
2024	SN Computer Science	Folino et al. (2024)	Small-BERT + Active Learning	PolitiFact (FakeNews Net)	0.88 ± 0.06	0.88 ± 0.05

2024	Frontiers in Big Data	Padalko et al. (2024)	Attention-based Bi-LSTM	WELFake	97.66	NR (test)
2025	current study	—	BERT + Threshold	Fake.csv, True.csv	98	0. 987



الشكل (٥): يظهر المخطط البياني لثلاث دراسات سابقة مقارنة مع دراستنا.

وبناءً على هذه النتائج، يُمكن اعتماد النموذج المقترح كأساس لتطوير أنظمة إعلامية ذكية أكثر كفاءة في مواجهة التضليل أثناء الأزمات.

المناقشة

في هذا البحث، تم بناء منظومة متكاملة تعتمد على نموذج BERT مطور، تم دمجه مع تقنيات إضافية مثل إدخال الضوضاء الاصطناعية والتحكم التفاعلي بالعتبة عبر واجهة Gradio ، بهدف تعزيز دقة كشف الأخبار الكاذبة في سياق الإعلام الأمني وإدارة الأزمات. وقد أظهرت

النتائج أن النموذج المحسّـــن تمكن من تحقيق أداء عالٍ تجاوز النماذج التقليدية، إذ بلغت دقة التصنيف ٩٨٪، مع قيمة F1 Score بلغت ٠.٩٨٧، وهي مؤشرات تفوق ملحوظ بالمقارنة مع النموذج السابق الذي استخدم BERT فقط وحقق دقة ٩٣٪ و F1 بلغت ٠.٩١.

يعكس هذا التحسـن الملحوظ في النتائج فعالية النهج الذي تم اتباعه في هذا البحث. فقد سـاهم إدخال نسبة ضوضاء اصطناعية بلغت ٥٪ من التسميات في تقوية النموذج ضد التحيز وزيادة قدرته على التعميم. كما أن تمكين المستخدم من التحكم بعتبة القرار (threshold) عبر واجهة Gradioالتفاعلية أضـاف بُعدًا عمليًا يسـمح بتكييف النموذج حسـب طبيعة البيانات وسـياق الاسـتخدام، مما يزيد من مرونته وقابليته للتطبيق في البيئات الواقعية. هذه المنهجية التفاعلية لم تكن متاحة في النماذج السابقة، مما يجعل هذه الدراسة أكثر واقعية واستجابة لمتطلبات الإعلام الأمني، خصـوصًـا في أوقات الأزمات إذ تكون المعلومات سـريعة التغير والتشـويش مرتفعًا. بالإضافة إلى ذلك، أظهرت الدراسة أن التدريب المستمر للنموذج عبر ثلاث حقب (epochs) سـاعد في تقليل الخسـارة وتحقيق اسـتقرار أعلى في الأداء مقارنة بالتدريب السـريع الذي اعتمد عليه النموذج السـابق. وتم تأكيد فعالية هذا النهج من خلال مقارنة النتائج مع دراسـات حديثة منشـورة في مجلات محكمة، إذ أن هذه الدراسـات لم توفّر إمكانيات التفاعل مع النموذج أو التحكم بالعتبة ولم تختبر قابلية التعميم في بيئة تحاكي الأزمات الأمنية.

عند مقارنة دراسـات بهذه الدراسـات، يظهر النموذج المقترح في بحثنا ليس فقط كمنافس قوي من حيث الأرقام، بل أيضًـا كحل عملي متكامل يلبي حاجات الإعلام الأمني من حيث الدقة، التفاعلية، والقدرة على مواجهة التشـويش المعلوماتي. ومن الجدير بالذكر أن دمج البيانات النصية من مصـادر موسـعة (أكثر من ٢١,٠٠٠ نص) قد لعب دورًا محوريًا في تعزيز موثوقية

النموذج وقدرته على كشف أنماط التضليل. كل هذه العناصر تجعل من البحث مساهمة علمية رصينة، ذات تطبيقات مباشرة في البيئة الأمنية والإعلامية المعاصرة، وتمهّد الطريق نحو بناء أنظمة إعلامية مدعومة بالذكاء الاصطناعي يمكن الوثوق بها في فترات الأزمات.

الاستنتاجات

أظهرت نتائج البحث أن نموذج BERT المحسّن، والمزود بضوضاء اصطناعية وواجهة تفاعلية (Gradio) ، بالاعتماد على نصوص مكتوبة. حقق النموذج دقة عالية بلغت ٩٨ % و F1 Score وصل إلى 0.987، متفوقًا على النماذج السابقة التي اعتمدت فقط على BERT التقليدي. دمج الضوضاء ساعد على تحسين التعميم، وتوسيع قاعدة البيانات إلى أكثر من ٢١,٠٠٠ نص زاد من قوة النموذج. كما أثبتت الواجهة التفاعلية فائدتها في ضبط العتبة (Threshold)حسب الحاجة، ما يوفر أداة ذكية وفعالة لدعم الإعلام الأمني في أوقات الأزمات، مع إمكانيات تطبيق واقعية ضمن بيئات حساسة. يتطلب النموذج التدريب العالي والمستمر على قواعد بيانات أكثر شمولية ليعطي نتائج فائقة الدقة بزمن حقيقي عالي السرعة.

التوصيات

اهم التوصيات التي من الجدير بالإشارة ذكرها في أدناه:

١. تعزيز استخدام النموذج المطور في المؤسسات الأمنية لمواجهة التضليل الإعلامي في الأزمات.
٢. ربط النظام بمنصات إعلامية رسمية للحصول على بيانات آنية يمكن التحقق منها.
٣. اعتماد التدريب التفاعلي وضبط العتبة كجزء من أدوات التحقق اليومية.

٤. تطوير النموذج باستخدام تقنيات أكثر تقدمًا مثل RoBERTa أو DistilBERT للحصول على أداء أعلى.

٥. توسيع قاعدة البيانات لتشمل مصادر متعددة اللغات وأنواع إعلامية مختلفة (مرئية ومقروءة).

٦. أن تقنيات الاتصالات المستخدمة في إدارة الأزمات ترتكز على مفاهيم فيزيائية مثل الانتشار الكهرومغناطيسي والبصريات، مما يعزز من قابلية تكامل هذه الدراسة في منظومات تعتمد على مبادئ الفيزياء لتأمين الاتصال وتحسين كفاءة تحليل البيانات تحت الضغط.

٧. النماذج المستخدمة لمعالجة النصوص وتصنيف الأخبار المضللة يمكن إعادة توظيفها لتحليل البيانات الناتجة عن أجهزة تعمل بمبادئ فيزيائية، مثل الحساسات الضوئية والكاميرات الطيفية في بيئات الأزمات.

المراجع:

1. Ahmed, A., & Khan, M. N. (2020). Comparative study of fake news detection techniques. International Journal of Computer Applications, 182(40), 1–6.
2. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the NAACL-HLT Conference.
3. Folino, G., Sabatino, A., & Pizzuti, A. (2024). Improving fake news detection through Small-BERT models and active learning. SN Computer Science, 5(2), 1–13. https://doi.org/10.1007/s42979-024-02511-7
4. Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1).
5. Jain, D. K., Yadav, R. P., & Tiwari, M. M. (2022). Fake news detection using deep learning models: A comparative study. In Proceedings of the IEEE ICCCNT Conference.
6. Kumar, A., & Gupta, S. (2021). Fake news detection using NLP techniques: A systematic review. IEEE Access, 9, 23420–23445.
7. Kumar, A., & Sebastian, M. A. (2021). Fake news detection using machine learning approaches: A systematic review. IEEE Access, 9, 9040–9077.
8. Mihalcea, R., & Radev, D. (2011). Graph-based natural language processing and information retrieval. Cambridge University Press.
9. Nandini, M., & Krishnamurthi, R. (2022a). Robustness of NLP models against noisy text. Information Processing & Management, 59(2).

10. Nandini, M., & Krishnamurthi, R. (2022b). Text noise and its effect on deep learning models: A case study. Information Processing & Management, 59(1).
11. Padalko, P., Shvets, A., & Petrov, I. (2024). Fake news detection using attention-based Bi-LSTM networks. Frontiers in Big Data, 7, 1–10. https://doi.org/10.3389/fdata.2024.1411590
12. Raza, M., Khan, A., Yousaf, M. H., & Hussain, S. (2025). Fake news detection using progressive BERT-based transfer learning. PLOS ONE, 20(3), 1–17. https://doi.org/10.1371/journal.pone.0312345
13. Shu, Z., Wang, Y., & Liu, M. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations, 19(1), 22–36.
14. Singh, B., & Sharma, M. (2022). BERT-based deep learning framework for detection of fake news. Expert Systems with Applications, 193.
15. Spence, D. R. (2019). Crisis communication, image repair, and reputation. Journal of Contingencies and Crisis Management, 26(1), 78–86.
16. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151.
17. Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework. Council of Europe.
18. Wolf, T., Debut, L., Sanh, V., et al. (2020). Transformers: State-of-the-art natural language processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP 2020), 38–45.