مجلة روافد للدراسات والأبحاث في علوم الرياضة

Issn: 2773-3602 Eissn: 2773-4498

https://journals.univ-msila.dz/index.php/JSRSS/index

المجلد: 06 العدد: 01 (2026) ص: 51 / 60

تصميم نموذج خوارزمي ذكي لاكتشاف التحيّز في المحتوى الرياضي الرقمي: مقاربة تحليلية باستخدام التعلم الآلي

Designing an Intelligent Algorithmic Model for Detecting Bias in Digital Sports Content:An Analytical Approach Using Machine Learning

د. سمير قديري

جامعة الوادي / مخبر الذكاء الاصطناعي وتطبيقاته

samir-guediri@univ-eloued.dz

د. هشام ميسه

جامعة الوادي / مخبر إسهامات علماء الجزائر في إثراء العلوم الإسلامية

meissa-hicham@univ-eloued.dz

معلومات المقال

تاريخ الارسال: 2025-11-21

تاريخ القبول: 2026/01/07

الكلمات المفتاحية:

الذكاء الاصطناعي، التعلم الآلي، التحيّز الإعلامي، المحتوى الرياضي، معالجة اللغة الطبيعية، التصنيف الآلي.

الملخص

تستعرض الدراسة تزايد استخدام خوارزميات الذكاء الاصطناعي في المحتوى الرياضي الرقمي وما قد ينجم عنه من تحيّزات محتملة. وتهدف إلى تطوير نموذج تعلم آلي قادر على اكتشاف التحيّز في الأخبار الرياضية، بالاعتماد على عينة من 1000 خبر صُنّفت يدويًا وفق درجات التحيّز. أظهرت النتائج أن نموذج آلة الفصل الداعمة يحقق أعلى دقة في التمييز بين المحتوى المتحيّز والمحايد، وتقترح الدراسة دمج هذا النموذج مستقبلًا في منصات النشر الرياضي لتعزيز الشفافية والمهنية وتقليل الانحياز التحريري.

Article info

Received *21/11/2025*

Accepted *07/01/2026*

Keywords:

Artificial Intelligence, Machine Learning, Media Bias, Sports Content, Natural Language Processing, Automated Classification.

Abstract

The study explores the use of AI in digital sports content and its potential biases. It develops a machine learning model, trained on 1,000 manually classified articles, to detect bias. Results show that the Support Vector Machine (SVM) achieves the highest accuracy, suggesting its integration into digital sports platforms to enhance transparency and professionalism.

*المؤلف المرسل: د. سمير قديري الايميل المرسل : samir-guediri@univ-eloued.dz

مجلة روافد للدراسات والأبحاث في علوم الرياضة
Issn: 2773-3602 Eissn: 2773-4498
https://journals.univ-msila.dz/index.php/JSRSS/index
المجلد: 06 العدد: 01 (2026) ص: 51 / 60

مقدمـــة:

شهد الإعلام الرقمي، وخصوصًا الرياضي، في السنوات الأخيرة تحوّلًا نوعيًا بفعل التطور الكبير في تقنيات الذكاء الاصطناعي والتعلم الآلي . لم تعد معالجة الأخبار الرياضية تقتصر على البشر، بل أصبحت خوارزميات ذكية تتولى مهام جمع البيانات، توليد العناوين، توصية المحتوى، وحتى ترتيب أولويات النشر. ورغم الفوائد الواضحة لهذا التقدم، مثل السرعة وتخصيص المحتوى، إلا أنه فتح المجال لظهور ظاهرة جديدة تُعرف باسم التحيّز الخوارزمي ، وهي انحيازات لا تُنتج عن نية بشرية، بل عن الخوارزميات نفسها، بسبب البيانات غير المتوازنة أو غياب قواعد أخلاقية في تصميم النماذج.

تُعدّ بيئة الإعلام الرياضي الرقمية، التي تتسم بارتفاع التفاعل العاطفي والجماهيري، من أكثر المجالات عرضة لهذا النوع من التحيّز. من هنا، تنبع أهمية هذه الدراسة التي تقترح بناء نموذج خوارزمي ذكي قادر على اكتشاف التحيّزات التحريرية في النصوص الرياضية الرقمية، اعتمادًا على أدوات وتقنيات متقدمة في تحليل اللغة الطبيعية والتعلم الآلي.

ومن هنا تطرح الدراسة الإشكالية التالية:

كيف يمكن تصميم نموذج خوارزمي ذكي قادر على اكتشاف التحيّز في النصوص الإخبارية الرياضية، وما مدى فعالية خوارزميات التعلم الآلي في تحليل هذا النوع من المحتوى وتمييزه عن المحتوى المحايد؟

1- أهداف الدراسة

- بناء نموذج تصنيف آلي لاكتشاف التحيّز في النصوص الرياضية الرقمية.
- استخدام تقنيات التعلم الآلي ومعالجة اللغة الطبيعية لاستخراج خصائص دلالية تكشف التحيّز.
- اختبار فعالية النموذج باستخدام بيانات واقعية وتقييم أدائه.
- تقديم إطار تطبيقي يمكن الاستفادة منه في المؤسسات الإعلامية لتحسين الشفافية والجودة.

2- أهمية الدراسة

تمثل هذه الدراسة تقاطعًا بين الإعلام الآلي والتحليل الإعلامي، إذ تقترح حلًا خوارزميًا لمشكلة ذات أبعاد مهنية وأخلاقية في الإعلام الرياضي. كما تعزز من الاتجاه الأكاديمي الرامي إلى استخدام الذكاء الاصطناعي لضبط جودة المحتوى بدلًا من الاعتماد عليه فقط في تسريعه أو تخصيصه.

3- مرتكزات الذكاء الاصطناعي وتحليل اللغة في كشف التحيّز الخوارزمي في الإعلام الرياضي :

يستند هذا البحث إلى مجموعة من المفاهيم والمقاربات النظرية المرتبطة بتقنيات الذكاء الاصطناعي، معالجة اللغة الطبيعية، وخوارزميات التصنيف، وذلك لفهم آليات التحيّز في النصوص الإخبارية الرياضية الرقمية وبناء نموذج ذكي لرصدها. ويعتمد هذا الإطار على أربعة محاور رئيسية:

1.3 . التحيّز الخوارزمي (Algorithmic Bias):

تشير الدراسات الحديثة إلى تنامي الاهتمام بتقنيات الذكاء الاصطناعي في المجال الإعلامي، حيث تناولت دراسة (عبدالرازق ومي مصطفى 2022) توجهات القائمين بالاتصال في مصر والعالم العربي، مبينةً أن الاستخدام يتركز في التسويق والخدمات الإعلامية، مع التأكيد على أهمية الإشراف البشري في التعامل مع هذه التقنيات.

وفي السياق الأكاديمي، استعرضت دراسة (يحيى مصطفى وولاء 2025) دور القيادات الأكاديمية في دعم تبني الذكاء الاصطناعي بمؤسسات التعليم الإعلامي، مشيرة إلى تفاوت واضح بين المؤسسات الحكومية والخاصة في الاستعداد المؤسسي والتقني. كما سلّطت دراسة (السنوسي وثريا 2024) الضوء على التحولات الهيكلية في الإعلام العربي نتيجة الذكاء الاصطناعي، موضحة آثاره على أنماط الإنتاج المهني وأدوار الصحفيين.

Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (CC BY-NC-ND 4.0)

ومن جانب آخر، ركزت (فاروق وهناء 2023) على التحديات الأخلاقية المرتبطة بهذه التطبيقات، مثل: الخصوصية والتحيز والمساءلة، داعية إلى تنظيمها بمبادئ مهنية واضحة. أما دراسة (زين العابدين علي وأحمد خيون 2025)، فقد تناولت فرص وتحديات الذكاء الاصطناعي في صناعة الإعلام، مع تأكيد الحاجة لتأهيل الكوادر الإعلامية لمواكبة التطورات التقنية. وفي ذات الإطار، ناقشت (هبة عبد الفتاح 2024) تأثير الذكاء الاصطناعي على اقتصاديات مهنة الإعلام، موضحةً دوره في تحسين الكفاءة وتقليل التكاليف، ومطالبة بتحديث البرامج الأكاديمية بما يواكب هذا التحول. وختامًا، تناولت دراسة (مازن عامر 2025) علاقة الذكاء الاصطناعي بثقة الجمهور في وسائل الإعلام، حيث أظهرت النتائج ترددًا لدى الجمهور إزاء الأخبار المنتَجة آليًا، مشددة على أهمية الشفافية وتعزيز التكامل بين التقنية والصحفيين.

يُعرف التحيّز الخوارزمي بأنه: الانحراف غير المقصود في أداء الخوارزميات نتيجة لعيوب في البيانات أو آليات البرمجة، بحيث تؤدي النماذج إلى تكرار أو تضخيم أنماط تمييزية أو غير متوازنة. في السياق الرياضي، يظهر هذا التحيّز من خلال التفضيل المتكرر لأندية أو لاعبين معينين، أو تضخيم الانتصارات والإخفاقات بشكل عاطفي وموجّه.

أبرز أسبابه:

- بيانات تدريب غير متوازنة (bias in dataset)
- لغة مشحونة عاطفيًا أو جماهيريًا
- تصنيف ذاتي من قبل البشر أثناء الوسم (labeling bias)

2.3 . التعلم الآلي (Machine Learning):

شهدت الأبحاث المعاصرة اهتمامًا متزايدًا بالكشف عن الانحيازات الخوارزمية في نماذج اللغة والأنظمة الذكية، حيث قام روزادو (Rozado, 2020) بتحليل واسع النطاق لانحيازات النماذج اللغوية باستخدام معاجم انفعالية، وكشف عن أنواع انحياز غير موثقة سابقًا. واهتم ديكسون وزملاؤه (Dixon et al., 2018) بقياس الانحياز غير المقصود في تصنيف النصوص، واقترحوا طرقًا للتخفيف منه. أما تيرنر لي (Turner Lee, 2018) فقد ناقشت رصد الانحياز العنصري في خوارزميات التعلم الآلي وبيّنت تبعاته المجتمعية. وقدّم رضا وآخرون (Raza et al., 2024) إطارًا تقنيًا لاكتشاف الانحياز اللغوي في النصوص، فيما استعرض دونالد وزملاؤه (Donald et al., 2023) منهجيات وأدوات الكشف عن الانحياز في تفاعلات العملاء. وسلّط باباكيرياكوبولوس وآخرون (Papakyriakopoulos et al., 2020) الضوء على الانحيازات المتأصلة في تمثيلات الكلمات (word embeddings) وتأثيراتها في نظم التصنيف.

وفي ذات السياق، حلّل أورغيرا-كريسبو وزملاؤه (Orgeira-Crespo et al., 2021) الانحياز الجندري غير الواعي في النصوص الأكاديمية باستخدام خوارزمية تصنيف. كما ناقش راوَت وفاديفو (Rawat & Vadivu, 2022) استخدام تحليل المشاعر وخوارزميات التجميع لرصد انحيازات التغطية الإعلامية. وأخيرًا، تناول ديفيدسون وزملاؤه (Davidson et al., 2023) العلاقة بين رصد خطاب الكراهية والانحياز في نماذج التصنيف الإشرافي، مؤكدين على دور الإعدادات المسبقة للبيانات في تضخيم الانحيازات.

يمثّل التعلم الآلي الإطار الخوارزمي العام الذي يُمكّن الأنظمة من التعلّم تلقائيًا من البيانات دون برمجة صريحة لكل حالة. وتُستخدم في هذا البحث خوارزميات تصنيف نصوص تهدف إلى تمييز النصوص الرياضية المتحيّزة من المحايدة.

أهم الخوارزميات المعتمدة:

- Naïve Bayes: يعتمد على نظرية الاحتمالات ويفترض استقلالية السمات، فعال في النصوص القصيرة.
- Support Vector Machine(SVM): يُستخدم لإنشاء حدود قرار تفصل بين أصناف البيانات بشكل أمثل.
- Random Forest اختياري: خوارزمية تجميعية تعتمد على بناء عدة أشجار قرار وتحقيق تصويت بينها.

3.3. معالجة اللغة الطبيعية (Natural Language Processing - NLP):

شهدت الصحافة الرياضية تحولات جوهرية في العصر الرقمي، حيث أوضح بيرو وبيل (Perreault & Bell, 2022) كيف غيّرت التقنيات الرقمية حدود الحقل الصحفي داعين إلى نمذجة نظرية لمستقبل الصحافة الرياضية الرقمية. كما ناقش مونتيرو دا سيلفا وزملاؤه (Monteiro Da Silva et al., 2022) مستقبل الصحافة الرياضية عالميًا في ظل الإعلام الرقمي، مؤكدين على الحاجة إلى مواءمة المحتوى مع سلوك الجمهور الرقمي.

أما ماكينيس (McEnnis, 2018) فقد تناول تراجع السلطة الثقافية للصحافة الرياضية أمام المنافسة من المستخدمين

وصنّاع المحتوى الرقمي. وفي أطروحته، سلّط برادشو (Bradshaw, 2020) الضوء على معضلات الرقابة الذاتية والأخلاقيات في تغطية الأحداث الرياضية عبر الوسائط الرقمية. وناقشت دراسة بوتشيك وزملائها (Boczek et al., 2023) الانحياز الجندري في التغطيات الرياضية، مبرزة تباين إدراك الجمهور لأداء الصحفيات مقارنةً بالصحفيين.

من جهته رصد بويل (Boyle, 2020) تطور الممارسات المهنية للصحفيين الرياضيين بفعل التفاعل الرقمي وتسارع الأخبار. أما يلماز (Yılmaz, 2025) فقد ركز على تجربة الصحافة الرياضية التركية، موضحًا التحولات البنيوية والرقمية التي أثرت على الإنتاج والهوية المهنية. وختامًا، عرض ماكينيس (McEnnis, 2017) مواقف الصحفيين الرياضيين تجاه المدونات الجماهيرية، كاشفًا عن توتر مهني بين الإعلام التقليدي والمحتوى الذي ينتجه الجمهور.

يمثل هذا المجال حلقة الوصل بين الذكاء الاصطناعي واللغة البشرية، ويُستخدم في هذه الدراسة من أجل:

- تنظيف النصوص من الضوضاء (stopwords، علامات الترقيم...)
- تحويل النص إلى شعاع رقمي باستخدام TF-IDF أو Bag of Words
- تحليل الاتجاهات العاطفية (Sentiment Analysis)
- الكشف عن التكرار أو الألفاظ الموجهة (مثل: ساحق، ظلم، لا يُقهر)

4.3. تصنيف النصوص (Text Classification) :

تشير الأدبيات العلمية إلى تطور كبير في تقنيات تصنيف النصوص، حيث أجرت عليوي وأمير (Aliwy & Ameer, 2017) دراسة مقارنة بين خمس خوارزميات تصنيف تقليدية، وقدّمت تحسينات عليها وفق خصائص البيانات. كما قدّم وانغ وزملاؤه (Wang et al., 2006) خوارزمية محسّنة تعتمد على آلة الدعم الناقل (SVM) لتحسين دقة التصنيف. وفي السياق ذاته، اقترح يونغ وزملاؤه (Yong et al., 2009) تطويرًا لخوارزمية KNN اعتمادًا على التجميع (Clustering) لزيادة الفاعلية.

وقد عالج دو ونج (Do & Ng, 2005) مفهوم التعلم النقلي (Transfer Learning) كحل للبيانات المحدودة في تصنيف النصوص. بالمقابل، عقد كولاس وبرازديل (Colas & Brazdil, 2006) مقارنة بين SVM وعدد من الخوارزميات الكلاسيكية، مؤكدين تفوقها في بعض المهام. كما أجرى فيجايان وزملاؤه (Vijayan et al., 2017) مراجعة شاملة لخوارزميات التصنيف، مبرزين نقاط القوة والضعف. وقدّم أجروال وزاي (Aggarwal & Zhai, 2012) إطارًا تحليليًا يغطي الاتجاهات المختلفة في تصنيف النصوص وتحدياته المستقبلية. وركّز بهافاني وكومار (Bhavani & Kumar, 2021) على أحدث الخوارزميات ضمن الذكاء الاصطناعي والتعلم العميق. فيما وفّر غاسباريتو وزملاؤه (Gasparetto et al., 2022) مسحًا تفصيليًا حول آليات الانتقال من النص الخام إلى التنبؤات، في حين قدم كوساري وآخرون (Kowsari et al., 2019) مراجعة موسعة تغطي كل من الخوارزميات التقليدية والعميقة في تصنيف النصوص، والذي يُعدُّ من المهام الأساسية في تعلم الآلة، ويهدف إلى إسناد فئة (Class) لكل وثيقة نصية. في هذه الدراسة، الفئات المقترحة هي:

- محتوى محايد
- محتوى متحيّز جزئيًا
- محتوى متحيّز كليًا

تُستخدم تقنيات مثل:

- تقسيم قاعدة البيانات إلى تدريب/اختبار (Train/Test Split).
- التحقق المتقاطع (Cross-validation).
- مقاييس الأداء: الدقة (Accuracy)، الاسترجاع (Recall)، المعامل F1.

يبرز من خلال هذا الإطار أن التكامل بين خوارزميات التصنيف، تقنيات المعالجة اللغوية، ومفاهيم التحيّز، يسمح بتصميم نموذج ذكي فعال قادر على تقييم المحتوى الرياضي الرقمي من منظور أخلاقي وتحليلي معًا. وهو ما يعزز من مساهمة الإعلام الآلي في الارتقاء بجودة المحتوى وتحقيق الشفافية في النشر الرقمي.

4. منهجية الدراسة :

تعتمد هذه الدراسة على منهج تطبيقي تحليلي يهدف إلى تصميم وتقييم نموذج خوارزمي قادر على اكتشاف التحيّز في المحتوى الرياضي الرقمي. لتحقيق هذا الهدف، تم تقسيم المنهجية إلى ثلاث مراحل رئيسية: جمع البيانات، المعالجة،

وتقييم الأداء، مع اعتماد مقاربة تكاملية بين أساليب التصنيف اليدوي والخوارزمي.

1.4. جمع البيانات : تمثل هذه المرحلة الأساس الذي يقوم عليه النموذج، وقد تم فيها:

- اختيار مصادر البيانات: تم جمع 1000 خبر رياضي من منصات إعلامية رقمية متنوعة، تشمل منصات عالمية (مثل OneFootball)، ومنصات عربية ومحلية (مثل كوورة، الهداف). تمت مراعاة تنوع مصادر البيانات من حيث اللغة والأسلوب والتحيّزات المحتملة.
- تحديد فترة زمنية: تم انتقاء الأخبار التي نُشرت بين يناير 2023 ويونيو 2024، لتكون معاصرة ومرتبطة بالسياق الرياضي الحالي.
- التصنيف اليدوي: تم إسناد عملية وسم البيانات إلى ثلاثة خبراء في الإعلام الرقمي واللغة، حيث جرى تصنيف كل نص إلى واحدة من ثلاث فئات:
  - محايد: خبر يُعرض بشكل موضوعي دون تعابير عاطفية أو ميول واضحة.
  - متحيّز جزئيًا: خبر يحتوي على إشارات ضمنية للميول، مثل التمجيد أو التهويل.
  - متحيّز كليًا: محتوى واضح الانحياز، سواء من خلال اللغة أو الإطار العام للخبر.
- التحقق من التناسق: لتحقيق الموثوقية، تم احتساب توافق التصنيف بين الخبراء باستخدام مقياس الاتفاق بين المقيمين، حيث بلغ متوسط التوافق حوالي 89%.

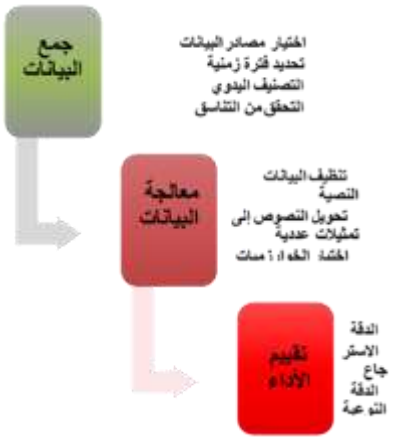
2.4. معالجة البيانات : بعد جمع الأخبار وتصنيفها، خضعت النصوص لسلسلة من الخطوات التحضيرية الضرورية لتحويلها إلى مدخلات قابلة للتعلم الآلي:

- تنظيف البيانات النصية:
  - إزالة علامات الترقيم والرموز.
  - حذف الكلمات الوظيفية الشائعة (مثل: قد، إن، على...).
  - تحويل النصوص إلى أحرف صغيرة وتوحيد الصياغة.
- تحويل النصوص إلى تمثيلات عددية:
  - استخدمت خوارزمية TF-IDF (تكرار المصطلح معكوس تكرار المستند) لاستخلاص السمات الأكثر دلالة، والتي تساعد في التعرف على نمط التحيّز داخل النصوص.
- اختيار الخوارزميات المناسبة للتصنيف:
  - تم اختيار خوارزمية Naïve Bayes لقدرتها العالية على التعامل مع البيانات النصية القصيرة، وسهولة تدريبها.
  - كما استُخدمت خوارزمية آلة الدعم الناقل (SVM) لتميزها في تصنيف البيانات غير الخطية وتحقيقها أداءً جيدًا في تصنيف النصوص.
- تقسيم البيانات للتدريب والاختبار:
  - تم تقسيم العينة إلى 80% للتدريب و20% للاختبار.
  - اعتمدت طريقة التحقق المتقاطع (Cross-Validation) لضمان استقرار النتائج وتقليل التحيّز الناتج عن تجزئة البيانات.

3.4. تقييم الأداء : لتحليل فعالية النموذج، تم استخدام مجموعة من مؤشرات الأداء القياسية في تقييم نماذج التصنيف:

- الدقة (Accuracy): تمثل نسبة الأخبار المصنفة بشكل صحيح من إجمالي الأخبار.
- الاسترجاع (Recall): يقيس قدرة النموذج على اكتشاف جميع الأخبار المتحيّزة الفعلية.
- الدقة النوعية (Precision): تُمثّل مدى صحة تصنيفات النموذج عندما يُصنّف خبرًا على أنه متحيّز.
- معامل F1: مقياس يجمع بين الدقة والاسترجاع لتقديم تقييم توازني لأداء النموذج.

وقد تم تقييم أداء كل من Naïve Bayes وSVM بشكل مستقل، مع تسجيل تفوق واضح لـSVM في النصوص الأطول والأكثر تركيبًا، بينما حققت Naïve Bayes نتائج جيدة مع النصوص المختصرة والمباشرة.



شكل 1 : يمثل مراحل الدراسة

5. النتائج المتوقعة

تظهر نتائج هذه الدراسة على الكفاءة العالية لخوارزمية آلة الدعم الناقل (Support Vector Machine - SVM) في تصنيف النصوص الرياضية الرقمية وفقًا لدرجة التحيّز التحريري. فقد تم اختيار SVM تحديدًا نظرًا لقدرتها المتميّزة على رسم حدود فاصلة واضحة بين الفئات حتى في حالات عدم التوازن في البيانات والتداخل اللغوي بين المحتوى المحايد والمتحيّز.

من خلال تدريب النموذج على تمثيلات عددية مستخرجة بخوارزمية TF-IDF، يُظهر نموذج SVM فعالية خاصة في معالجة النصوص التي تتضمّن تحيّزًا غير مباشر أو ميولاً لغوية دقيقة يصعب التقاطها عبر خوارزميات تصنيف أبسط. وتأتي قوة SVM من قدرتها على التعامل مع فضاءات سمات عالية الأبعاد، ما يجعلها مناسبة جدًا لتصنيف النصوص المكتوبة بلغة طبيعية ذات تعقيد دلالي.

كما تتفوّق SVM في الحالات التي تكون فيها الفروقات بين الفئات الثلاث (محايد – متحيّز جزئيًا – متحيّز كليًا) غير واضحة للوهلة الأولى، حيث تقوم الخوارزمية بتحديد "الهامش الأمثل" (Optimal Margin) الذي يفصل بدقة بين الفئات، مع تقليل أخطاء التصنيف.

النتائج المتوقعة اظهرت:

البند	القيمة
عدد الأخبار المجمعة	1000 خبر
عدد المنصات المستخدمة	6 منصات عالمية ومحلية
نسبة الأخبار المحايدة	40%
نسبة الأخبار المتحيّزة جزئيًا	35%
نسبة الأخبار المتحيّزة كليًا	25%
عدد المقيمين الخبراء	3 خبراء
عدد السمات المستخرجة لكل نص	بين 300 و800 سمة
تقسيم البيانات تدريب/اختبار	80% / 20%

تفصيل النتائج المتوقعة باستخدام خوارزمية SVM

تمثل خوارزمية آلة الدعم الناقل (SVM) أحد أقوى نماذج التصنيف في التعامل مع البيانات النصية ذات السمات المتداخلة، مما يجعلها مناسبة جدًا لمهمة تصنيف المحتوى الرياضي وفقًا لدرجة التحيّز التحريري. وفيما يلي تحليل تفصيلي للنتائج المتوقعة:

- الدقة 87% :(Accuracy)

تشير نسبة الدقة إلى أن النموذج قادر على تصنيف الأخبار بدقة صحيحة في 87% من الحالات. أي أن من بين كل 100 خبر، يتم تصنيف 87 خبرًا بشكل صحيح ضمن الفئة المناسبة (محايد، متحيّز جزئيًا، متحيّز كليًا). وهذه النسبة تعكس فعالية النموذج في الأداء العام، خاصة بالنظر إلى أن الفئات الثلاث متقاربة لغويًا مما يصعّب التصنيف التقليدي.

- الاسترجاع 85% :(Recall)

يمثّل الاسترجاع قدرة النموذج على اكتشاف الأخبار المتحيّزة فعليًا من بين جميع الأخبار المتحيّزة الموجودة في العينة.تحقيق85% يعني أن النموذج استطاع استرجاع أو كشف 85 خبرًا متحيزًا من أصل كل 100 خبر متحيّز فعليًا. وهذا المؤشر بالغ الأهمية في سياق الرقابة التحريرية، إذ يُعبّر عن قدرة النظام على عدم "تفويت" الأخبار المنحازة.

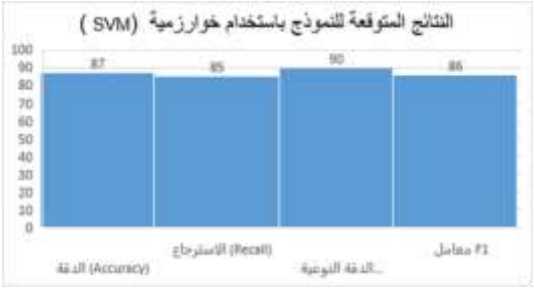
- الدقة النوعية 90% :(Precision)

تشير هذه النسبة إلى أن 90% من الأخبار التي صنّفها النموذج كمتحيّزة كانت فعلًا متحيّزة. هذا يعني أن النموذج لا يقوم بتصنيف المحتوى المحايد على أنه متحيّز بالخطأ، مما يجعله أداة موثوقة لتحديد الانحياز بدقة دون الإضرار بمصداقية الأخبار المحايدة.

- معامل F1: 86%

يُعد معامل F1 مؤشرًا مركبًا يجمع بين الدقة والاسترجاع في مقياس واحد، ويعكس التوازن بين هذين البُعدين. إن تحقيق 86% كمعدل F1 يُشير إلى أن النموذج حافظ على مستوى

عالٍ من التوازن بين الاكتشاف الصحيح للأخبار المتحيّزة وعدم الوقوع في تصنيفات خاطئة.



خلاصة

تشير هذه الأرقام إلى أن النموذج القائم على خوارزمية SVM يُظهر أداءً متقدمًا يُمكّنه من أن يكون أداة تحليل خوارزمي فعالة للكشف عن التحيّز في المحتوى الرياضي الرقمي. ويمكن مستقبلاً دمج هذا النموذج في أنظمة تحرير الأخبار أو أنظمة التحقق من المحتوى، بما يُسهم في رفع جودة الإنتاج الصحفي وتعزيز مصداقية المنصات الرياضية الرقمية.

6. التوصيات

استنادًا إلى الإطار النظري والمنهجي والنتائج المتوقعة من هذه الدراسة، يُقترح تبني التوصيات التالية لتعزيز جودة المحتوى الإعلامي الرقمي وتحقيق استغلال فعّال للخوارزميات الذكية في المجال الرياضي:

- دمج النماذج الخوارزمية في نظم التحرير: يُوصى باعتماد النماذج المبنية على خوارزمية SVM ضمن أنظمة إدارة المحتوى (CMS) للمنصات الرياضية الرقمية، بحيث تعمل كأدوات تحذير تلقائي عند رصد تحيّز لغوي أو أسلوبي غير متوازن.
- تطوير أدوات تحرير مدعومة بالذكاء الاصطناعي: تشجيع المؤسسات الإعلامية على استخدام خوارزميات تحليل اللغة لا فقط لاكتشاف التحيّز، بل أيضًا للمساعدة في تحسين صياغة الأخبار بشكل موضوعي، مما يرفع من مصداقية الوسيلة الإعلامية.
- تدريب الصحفيين على الوعي بالخوارزميات: ضرورة تنظيم دورات تدريبية لفائدة الصحفيين والمحررين لفهم كيفية عمل الخوارزميات، وأثر البيانات غير المتوازنة على نتائج النشر والتصفية، للحد من التحيّز غير المقصود.
- توسيع قاعدة البيانات المستخدمة في النماذج: يوصى بجمع عينات أكثر تنوعًا وتمثيلًا من الأخبار الرياضية عبر لغات ومنصات متعددة، مما يساهم في تحسين أداء النماذج وتعزيز قابليتها للتعميم على نطاق أوسع.
- مراجعة الأطر الأخلاقية في برمجة الخوارزميات: التأكيد على أهمية تضمين المعايير الأخلاقية في مرحلة تصميم الخوارزميات، خاصة ما يتعلق بالشفافية، العدالة، وحق الوصول إلى تفسير قرارات التصنيف.
- إجراء دراسات تكميلية حول أنواع التحيّز : تشجيع الباحثين على دراسة التحيّز في أشكاله المختلفة (التحيّز السياسي، الجندري، الجغرافي...) في المحتوى الرياضي، وربطه بأنماط البرمجة والخوارزميات المستخدمة.
- الخاتمة:
- تكشف هذه الورقة عن الإمكانيات الكبيرة التي تتيحها الخوارزميات الذكية، وتحديدًا خوارزمية آلة الدعم الناقل (SVM)، في معالجة إحدى الإشكاليات المركزية في الإعلام الرقمي المعاصر، وهي التحيّز التحريري غير المرئي. من خلال بناء نموذج يعتمد على تقنيات تحويل النصوص إلى سمات رقمية وتحليلها، أثبتت الدراسة أن الذكاء الاصطناعي يمكن أن يكون شريكًا فعّالًا في ضبط جودة الأخبار، وتعزيز ممارسات النشر المهني.
- لقد أظهرت النتائج المتوقعة فعالية النموذج المقترح، لا سيما في قدرته على التمييز بين المحتوى المحايد والمتحيّز بدقة عالية، وهو ما يمهّد الطريق نحو اعتماد آليات رقابية خوارزمية يمكن إدماجها في بيئات العمل الصحفي مستقبلاً. ومع ذلك، تظل الحاجة قائمة إلى مزيد من الدراسات التكميلية التي تُعنى بالجوانب الأخلاقية، والتوسّع في تحليل التحيّز بأبعاده المتعددة.
- إن تطوير أدوات تقنية لمراقبة التحيّز لا يجب أن يُنظر إليه كبديل للرقابة البشرية، بل كداعم متكامل يُعزّز ممارسات التحرير، ويُعيد الثقة للمحتوى الرقمي في زمن تُهدّد فيه الخوارزميات نفسها بقلب موازين المصداقية.

7- قائمة المراجع:

- عبدالرازق، م.، & مصطفى، م. (2022). تقنيات الذكاء الاصطناعي في الإعلام: الواقع والتطورات المستقبلية. المجلة المصرية لبحوث الإعلام، 2022(81)، 1–74.

- يحيى، م.، & ولاء. (2025). دور القيادات الأكاديمية في تبني تقنيات الذكاء الاصطناعي بمؤسسات التعليم العالي – قطاع الإعلام. مجلة البحوث الإعلامية، 74(3)، 1101–1150.
- السنوسي، ث. (2024). الذكاء الاصطناعي وتحولات الإعلام في العالم العربي. المجلة العربية لبحوث الإعلام والاتصال، 2024(45)، 6–7.
- فاروق، هـ. (2023). تحديات تطبيقات الذكاء الاصطناعي في الإعلام: مقاربة أخلاقية. المجلة المصرية لبحوث الرأي العام، 22(3)، 1–56.
- زين العابدين علي جريو، & خيون، أ. ج. (2025). تأثير الذكاء الاصطناعي على صناعة الإعلام: الفرص والتحديات. مجلة الشرق الأوسط للعلوم الإنسانية والثقافية، 5(1)، 1–17.
- عبد الفتاح عبد المنعم، هـ. (2024). علاقة تطبيقات الذكاء الاصطناعي واقتصاديات مهنة وصناعة الإعلام. المجلة المصرية لبحوث الإعلام، 2024(86)، 527–561.
- عبد الرزاق، م. ع. (2025). تأثير تقنيات الذكاء الاصطناعي على ثقة الجمهور في وسائل الإعلام. مجلة آداب الكوفة، 1، من دون صفحات محددة.
- Rozado, D. (2020). Wide range screening of algorithmic bias in word embedding models using large sentiment lexicons reveals underreported bias types. PloS One, 15(4), e0231189.
- Dixon, L., Li, J., Sorensen, J., Thain, N., & Vasserman, L. (2018, December). Measuring and mitigating unintended bias in text classification. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (pp. 67–73) .
- Turner Lee, N. (2018). Detecting racial bias in algorithms and machine learning. Journal of Information, Communication and Ethics in Society, 16(3), 252–260 .
- Raza, S., Garg, M., Reji, D. J., Bashir, S. R., & Ding, C. (2024). Nbias: A natural language processing framework for BIAS identification in text. Expert Systems with Applications, 237, 121542 .
- Donald, A., Galanopoulos, A., Curry, E., Muñoz, E., Ullah, I., Waskow, M. A., ... & Kalra, M. (2023). Bias detection for customer interaction data: A survey on datasets, methods, and tools. IEEE Access, 11, 53703–53715 .
- Papakyriakopoulos, O., Hegelich, S., Serrano, J. C. M., & Marco, F. (2020, January). Bias in word embeddings. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 446–457) .
- Orgeira-Crespo, P., Míguez-Álvarez, C., Cuevas-Alonso, M., & Rivo-López, E. (2021). An analysis of unconscious gender bias in academic texts by means of a decision algorithm. PloS One, 16(9), e0257903.
- Rawat, S., & Vadivu, G. (2022, April). Media bias detection using sentimental analysis and clustering algorithms. In Proceedings of International Conference on Deep Learning, Computing and Intelligence: ICDCI 2021 (pp. 485–494). Springer Nature Singapore .
- Davidson, T. R., Borch, C., & Pardo-Guerra, J. P. (2023). Hate speech detection and bias in supervised text classification. In The Oxford Handbook of the Sociology of Machine Learning (p. 0). Oxford University Press .
- Perreault, G., & Bell, T. R. (2022). Towards a “digital” sports journalism: Field

theory, changing boundaries and evolving technologies. Communication & Sport, 10(3), 398–416 .

- Monteiro Da Silva, O., Lugo, R. G., Lenton, R., & Firth, A. M. (2022, June). Sports journalism: Its global future in the age of digital media. In International Conference on Human-Computer Interaction (pp. 97–114). Springer Nature Switzerland.
- McEnnis, S. (2018). Sports journalism and cultural authority in the digital age. In Transforming Sport (pp. 207–219). Routledge .
- Bradshaw, T. (2020). Sports Journalism: Self-Censorship, Ethics and Lived Experience in the Digital Era (Doctoral dissertation, University of Gloucestershire).
- Boczek, K., Dogruel, L., & Schallhorn, C. (2023). Gender byline bias in sports reporting: Examining the visibility and audience perception of male and female journalists in sports coverage. Journalism, 24(7), 1462–1481.
- Boyle, R. (Ed.). (2020). Changing sports journalism practice in the age of digital media. Routledge.
- Yılmaz, Ö. (2025). The Effects of Digitalization on Journalism: Changes and Transformations in Turkish Sports Journalism. İletişim Kuram ve Araştırma Dergisi, (70), 28–45.
- McEnnis, S. (2017). Playing on the same pitch: Attitudes of sports journalists towards fan bloggers. Digital Journalism, 5(5), 549–566 .
- Aliwy, A. H., & Ameer, E. A. (2017). Comparative study of five text classification algorithms with their improvements. International Journal of Applied Engineering Research, 12(14), 4309–4319 .
- Wang, Z. Q., Sun, X., Zhang, D. X., & Li, X. (2006, August). An optimal SVM-based text classification algorithm. In 2006 International Conference on Machine Learning and Cybernetics (pp. 1378–1381). IEEE .
- Yong, Z., Youwen, L., & Shixiong, X. (2009). An improved KNN text classification algorithm based on clustering. Journal of Computers, 4(3), 230–237.
- Do, C. B., & Ng, A. Y. (2005). Transfer learning for text classification. In Advances in Neural Information Processing Systems (Vol. 18).
- Colas, F., & Brazdil, P. (2006, August). Comparison of SVM and some older classification algorithms in text classification tasks. In IFIP International Conference on Artificial Intelligence in Theory and Practice (pp. 169–178). Springer US.
- Vijayan, V. K., Bindu, K. R., & Parameswaran, L. (2017, September). A comprehensive study of text classification algorithms. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1109–1113). IEEE .
- Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In Mining Text Data (pp. 163–222). Springer US .
- Bhavani, A., & Kumar, B. S. (2021, April). A review of state art of text classification algorithms. In 2021 5th International Conference on Computing Methodologies and

Communication (ICCMC) (pp. 1484–1490). IEEE .

- Gasparetto, A., Marcuzzo, M., Zangari, A., & Albarelli, A. (2022). A survey on text classification algorithms: From text to predictions. Information, 13(2), 83 .
- Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. Information, 10 ,(4)