منظم من طرف: المركز الجامعي سي الحواس _بريكة _مخبر آفاق الحوكمة للتنمية المحلية المستدامة_

25 سبتمبر 2022

"التحيز الرقمي في أنظمة الذكاء الاصطناعي"

"Digital Bias In Intelligence Systems"

المحور الثالث: تأثيرات استخدام أنظمة الذكاء الاصطناعي على الحقوق والحريات

1. بلقيس قرازة طالبة دكتوراه ل.م.د. 2. د. سعاد قصعة، أستاذ محاضر_ أ_

Belkis Grazza D. Souad Kassaa

1جامعة الأمير عبد القادر، كلية الشريعة والاقتصاد، قسم: الشريعة والقانون، belkisbilla@gmail.com

2جامعة الأمير عبد القادر،كلية الشريعة والاقتصاد، قسم الشريعة والقانون، habibkassa@gmail.com

الملخص

رغم التقدم الذي حققه الذكاء الاصطناعي في مختلف المجالات، إلا أن النتائج التي كانت تحققها أنظمته لم تكن تتسم بالحيادية بل بالعنصرية فأصبح لدينا ما يسمى اليوم بالتحيز الرقمي أو العنصرية الرقمية أو التحيز الخوارزمي، لهذا كان هدف الدراسة تسليط الضوء على أسبابه مع أن السبب الرئيس هم البشر أنفسهم، فنظام التعلم الآلي لخوارزميات الذكاء الاصطناعي يعتمد على البيانات القاعدية التي غالبا ما تكون مشبعة بتحيزات البشر وعنصريتهم، وعلى الآليات والمبادرات المتخذة من طرف المنظمات والشركات الكبرى للتخفيف من حدة المشكلة، للوصول في الأخير إلى حقيقة أن التحيز الرقمي ما هو إلا امتداد للتحيز البشري وهذا لارتباط التعلم الآلي لهذه الأنظمة بالبيانات المتوافر أغلبها على الانترنيت، وأن معظم القرارات المتخذة للحد منه لم تقلل من هذا التحيز إلا بنسب قليلة.

الكلمات المفتاحية: التحيز الرقمي_ أنظمة_ الذكاء الاصطناعي.

Summary

Despite the progress made by artificial intelligence in various fields, the results achieved by its systems were not characterized by neutrality, but by racism, so we have what is called today digital bias, digital racism, or algorithmic bias. That is why the aim of the study was to shed light on its causes, although the main reason is humans themselves. The machine learning system for artificial intelligence algorithms relies on base data that is often saturated with human biases and racism, and on the mechanisms and initiatives taken by organizations and large companies to mitigate the problem, to reach Finally, to the fact that digital bias is only an extension of human

bias, and this is because the machine learning of these systems is linked to data, most of which are available on the Internet, and that most of the decisions taken to reduce it did not reduce this bias except by a few percentages.

Keywords: digital bias, systems, artificial intelligence

1. مقدمة

يعيش العالم اليوم تحديات راهنة في مجال التكنولوجيا وأتمتة القطاعات، خاصة بعد الثورة الصناعية الرابعة التي أضفت تحولا جذريا في مختلف المجالات، ومن نتاج هذه الثورة ما يسمى بالذكاء الصناعي، هذا الأخير أصبح المؤشر الأساسي على تقدم الدول وتطورها على مختلف الأصعدة مما دفع بالعديد منها إلى تبنيه بنسب متفاوتة.

إن هذا الاعتماد الشبه الكلي على أنظمة الذكاء الصناعي كان من شأنه أن يثير حفيظة رجال القانون خاصة في مسألة أخلقة التكنولوجيا، فكان لزاما على المبرمجين برمجة الأجهزة على بعض القيم الانسانية التي من شأنها أن تشكل بغيابها خطرا على حياة الانسان، ومن أهم هذه القيم الحيادية وعدم التحيز إذ السمة التي اصبحت تتصف بها هذه الأنظمة هي اللاحيادية والعنصرية في النتائج التي تقدمها محدثة بذلك ما يمكن تسميته بالتحيز الرقمي أو العنصرية الرقمية.

لقد برزت هذه المشكلة بعد أتمتة بعض القطاعات الصناعية والإدارية والأمنية واعتمادها على نظام الانتقاء بحيث اعتمدت ووظفت خوارزميات تصنف الأشخاص على أساس عدة معايير خاطئة مما يعني أن الذكاء الصناعي صار يمس بالحقوق الأساسية للفرد وهي حقه في المساواة مع غيره، ومنه يمكن القول أن **<u>اشكالية البحث</u>** كالآتي: **كيف تساهم أنظمة الذكاء الاصطناعي في تعزيز العنصرية بين الأفراد؟**

يندرج تحت هذه الإشكالية جملة من **<u>الأسئلة الفرعية</u>** التالية:

1) ما هو تعريف التحيز الرقمي؟
2) فيم تتمثل أسباب العنصرية الرقمية في أنظمة الذكاء الاصطناعي؟
3) ما هي الحلول التي وجدت على الساحة الدولية والتدابير الوقائية المتخذة لحل أزمة التحيز الرقمي للذكاء الاصطناعي؟

<u>الفرضيات:</u>

1) يرتبط التحيز الرقمي لأنظمة الذكاء الاصطناعي بالبيانات المنحازة والخوارزميات المعيبة.

2) إن كبرى الشركات التكنولوجية تسعى كلها لأخلقة الذكاء الاصطناعي بما يوافق حقوق الانسان وحرياته الاساسية.

أهداف الدراسة:

1) تسليط الضوء على الأسباب الرئيسية لانحياز خوازميات الذكاء الاصطناعي حتى يصبح معالجتها أمرا ممكنا.
2) محاولة استخلاص حلول جذرية للحد من العنصرية الرقمية، أو التخفيف من نتائجها الخاطئة.

منهج الدراسة:

تم الاعتماد على كل من المنهج الوصفي والمنهج الاستقرائي، حيث تم الاعتماد على الأول منهما في وصف مشكلة التحيز الرقمي وبيان أسبابها، بينما تم الاعتماد على الثاني في استقراء كل من نماذج التحيز وآليات الحد منه في الفترة الأخيرة.

خطة الدراسة:

1. مفهوم التحيز الرقمي في أنظمة الذكاء الاصطناعي.
2. أسباب التحيز الرقمي في أنظمة الذكاء الاصطناعي.
3. آليات الحد من تحيز أنظمة الذكاء الاصطناعي.

2. مفهوم التحيز الرقمي في أنظمة الذكاء الاصطناعي.

من مخرجات الثورة الصناعية الرابعة التي جعلت العالم يدخل حقبة تقنية غيرت مجراه هي أنظمة الذكاء الاصطناعي، هذه الأخيرة كغيرها من مفرزات التكنولوجيا لا يقتصر دورها على تقديم مختلف الخدمات التقنية لمستخدميها وتسهيل حياتهم فقط، بل يطول دورها إلى الإضرار بمصالح الأفراد وخاصة حرياتهم وحقوقهم الأساسية، فنجدها من هذا الجانب قد مست حقهم في المساواة والعدالة، وأصبحت هي الأخرى وسيلة للتمييز بينهم، وهذا ما يسمى حاليا بالتمييز أو العنصرية الرقمية، في هذا المحور سيتم التطرق إلى التعريف الاجرائي للتحيز الرقمي، بيان أسبابه، ثم إدراج جملة من الأمثلة الرائجة التي بها شاع التمييز الرقمي.

1.2 تعريف التحيز الرقمي

جاء معنى التحيز في معجم البيانات والذكاء الاصطناعي على أنه: انحياز نظام الذكاء الاصطناعي وتفضيله لمجموعات معينة على أخرى، ويطلق عليه أيضا "تحيز خوارزمي"،[1] ويمكن تعريفه أيضا على أنه "حالة انحراف في نتائج خوازميات[2] التعلم الآلي يحدث بسبب وجود فرضيات متحيزة أثناء عملية تطوير الخوارزمية وتمثل انعكاسا لعنصرية المجتمع وتحيزه ضد فئة معينة أو قد يكون نتيجة تحيز بيانات التدريب التي يتم تغذية نظام الذكاء الاصطناعي بها". [3]

إذا يمكن القول بأن التحيز الرقمي هو أن تنحاز الخوازميات المعتمدة في أنظمة الذكاء الاصطناعي لطرف معين فتعطي نتائج تتسم بالعنصرية والطبقية وهذا بناء على البيانات التي اعتمدت عليها في مرحلة التعلم الآلي.

2.2 أنواع التحيز الرقمي

إن العنصرية أو التحيز الذي قد تتسم به خوازميات الذكاء الصناعي على ثلاث أضرب هي كالآتي:

---

[1] معجم البيانات والذكاء الاصطناعي انجليزي عربي، الهيئة السعودية للبيانات والذ الاصطناعي SDAIA، السعودية، ط1، 2022م، ص 46.

[2] جمع خوارزمية وهي كلمة مشتقة من اسم عالم الرياضيات المسلم محمد بن موسى الخوارزمي وتعرف على أنها "مجموعة من الخطوات (التعليمات) المرتبة والمنتهية، لتنفيذ عملية حسابية أو منطقية، أو غيرها بشكل تتابعي متسلسل ومنطقي"، السيد بدر، تصميم وتحليل الخوازميات، المكتبة العربية، د ط، ص 6.

[3] "كيف يمكن التغلب على تحيز الذكاء الاصطناعي"، https://masaar.net، 2022/08/21، 22: 15.

1) **الانحياز التفاعلي:** وهو أحد أشكال الانحياز التي أُحدثت للذكاء الاصطناعي وهذا بسبب تفاعل المستخدمين معه، فمثلا في حال التطبيقات التي يرفع مستخدميها صورهم ويحولونها إلى البشرة البيضاء يكتسب التطبيق خبرة أكبر فيقوم بتصورات أقرب للمستخدمين ذوي البشرة البيضاء لأن غالبيتهم تفاعلوا معه.

2) **الانحياز الكامن:** يقوم الذكاء الاصطناعي غالبا بدراسة مجموعة من البيانات لإيجاد الروابط بينها والتعلم منها وغالب الوقت تكون تلك البيانات منحازة إلى فئة على حساب أخرى، مثلا اللغة الانجليزية تحتل 55 % من الانترنيت، بينما العربية تحتل 1.6% فقط، وعليه فإن ترجمة غوغل ستقوم بترجمة اللغة الانجليزية أفضل من ترجمتها للعربية.

3) **الانحياز الاختياري:** وهو الانحياز في شكل البيانات التي يختارها المبرمج لتدريب الذكاء الاصطناعي عليها، فمثلا إذا أراد تدريب النظام على تمييز الجنس هل رجل أم امرأة واعتمد على الصور الشخصية في ملفات تويتر فإنّ النتائج ستُظهر الولايات المتحدة الأمريكية في الصدارة لأنها تتصدر عدد مستخدمي تويتر، وعليه فإن اختيار المبرمج لهذه البيانات بالتحديد خلق انحيازا في الخوارزمية.[1]

3.2 نماذج عن أنظمة متحيزة

كما طال الذكاء الاصطناعي جل القطاعات فكذلك فعل التحيز التابع له، فلا نجد نظاما إلا وفيه احتمالية نشوب التحيز فيه، ولتعدد الأمثلة والنماذج ارتأينا التركيز بالشرح الوافي على نظامين اثنين رائجين عُرفت على مستواهم نسبة كبيرة من العنصرية الرقمية وهي أنظمة التعرف على ملامح الوجه، وأنظمة التنبؤ الإجرامي، واختص هذان النموذجان دون غيرهما كون الأول منهما جد عام إذ فيه تندرج أهم صور العنصرية وهي العنصرية على أساس اللون والجنس، والثاني لأنه نظام جد خطير يمس بأمن الفرد وخصوصيته ويمس بالعدالة فكانا كلاهما مناسبا للإدراج، على أن يتم إرفاق هذين النظامين بنماذج أخرى.

1) أنظمة التعرف على ملامح الوجه

"يعتبر تمييز الوجه أحد أكثر التطبيقات التي لها علاقة بتحليل الصورة، حيث يعتبر تحدي حقيقي لبناء نظام آلي يعادل قدرة الإنسان على تمييز الوجه، على الرغم من قدرة الإنسان على تحديد الوجوه

[1] مصطفى السيد حسين، "عنصرية الذكاء الصناعي... لماذا يبدو أصدقاؤك بملامح أوروبية عند استخدامهم تطبيقات وفلاتر الذكاء الاصطناعي؟"، www.arabicpost.net، نشر بتاريخ 2020/07/08، تاريخ الاطلاع: 2022/08/22، 14:35.

المعروفة،.ولكنه ليس قادرا على التمييز عندما يكون عليه التعامل مع كمية كبيرة من الوجوه الغير معروفة"، [1] لهذا نجد في وقتنا الحالي تنافس كبير بين الشركات المصنعة لهذه التقنية، ولعل من أبرز المطورين لها شركات (مايكروسوفتMICROSOFT) و(أمازون AMAON) و(آي بي أم IBM) وكلها شركات أمريكية، وتتصدر الصين قائمة الدول المستخدمة لهذه التقنية وتليها الولايات المتحدة الأمريكية،[2] ورغم الايجابيات المرجوة من هكذا أنظمة إلا أن النتائج السلبية التي أدت إليها الماسة بكرامة الأفراد وحقهم في المساواة بغض النظر عن العرق والجنس واللون والثقافة أثارت حفيظة المستخدمين والمسؤولين مما جعل الضوء يسلط عليها في ما إذا كانت حقا تؤدي إلى تعزيز التحيز بين الأفراد، في هذا الصدد كشفت عدة دراسات أن برامج التعرف على ملامح الوجه والصور هي برامج تنطوي على عنصرية واضحة وعلنية، فهي تعطي نتائج متحيزة إما على أساس اللون أو الجنس، من هذه الدراسات تلك التي أجرتها الحكومة الاتحادية في الولايات المتحدة الأمريكية وهي دراسة حديثة في هذا الخصوص والتي توصلت إلى أن أنظمة التعرف على الوجه المعدة في المطارات والموانئ والجمارك ووكالات انفاذ القانون وكل مؤسسة تعتمد تلك التقنية مليئة بالأخطاء ذات الدلالة العنصرية، إذ بينت أن الأشخاص الأمريكيين من أصل افريقي أو آسيوي أكثر عرضة للتمييز بشكل خاطي 100 مرة من الرجال البيض، ورغم وجود عدة شركات تعمل على تطوير تقنيات التعرف على الوجه مثل مايكروسوفت (MICROSOFT) وباناسونيك (PANASONIC) وانتل (INTEL) اختارت الحكومة الأمريكية الخوازميات التي طورتها شركة أمازون (AMAZON) والتي أثبتت الدراسات تحيزها الواضح ضد السود.[3]

من الأمثلة أيضا واقعة تجربة الكاميرا التي طورتها احدى شركات الحاسوب حيث تعرفت الكاميرا على وجه واندا زامن ذات البشرة البيضاء وتتبع حركاتها، ولم تتعرف على وجه زميلتها ذات البشرة السوداء ديسي

[1] محمد مازن المصطفى، دمج خوارزميات التعرف على الوجه مع سرعة الطباعة من أجل المصادقة على الأنترنيت، رسالة ماجيستير، قسم الدراسات العليا علوم الويب، الجامعة الافتراضية السورية، سوريا، 2018/2017، ص 4.

[2] طه الراوي، "التعرف على الوجه كل ما تحب معرفته عن التقنية التي تحفظ ملامحنا"، www.noonpost.com نشر بتاريخ 2020/06/24، تاريخ الاطلاع 2022/08/24، 14:34.

[3] طه الراوي، "عنصرية التكنولوجيا...عن أسباب وأشكال التحيز الرقمي"، www.noonpost.com، نشر بتاريخ 2020/06/12، تاريخ الاطلاع 2022/08/21، 15:36.

كرير ولم يتتبع حركاتها كما قام برنامج beaty al الذي تم تطويره لعمل مسابقة جمال متنقلة حيث ابدى تحيزات عنصرية واضحة واعطى الاولية لذوي الشبرة البيضاء.[1]

إن مثل هذا التمييز الذي تعززه برامج التعرف على الوجه من شأنه المساس بمصداقية النظام بل بمصداقية الشركات الكبرى المؤسسة لها، فبينما كانت البشرية تنتظر اضافة جلية من أنظمة الذكاء الصناعي في مجال حقوق الإنسان وحرياته الأساسية بإزالة الفروق العرقية والثقافية والجنسية بين الشعوب والأفراد، أضحت هذه البرامج هدامة لهذه الآمال ومخيبة للظن لكثير من الحقوقيين فضلا عن من تم تحييدهم شخصيا أو تحييد الفئة التي ينتمون لها.

2) أنظمة التنبؤ الإجرامي

من مخرجات الذكاء الاصطناعي القدرة على التنبؤ بالجرائم قبل حدوثها وهذا باعتمادها على مكونات تقنية هي خوازميات، (ويمكن وصف تنبأ الجريمة بواسطتها بالتنبؤ الخوارزمي للجريمة) وهذا من خلال مفاهيم الإحصاء وتحليل البيانات الضخمة، واكتشاف الأنماط وربط بعضها ببعض واعطاء النتائج وبيانات ضخمة وهي معطيات يتم تغذية تطبيقات الذكاء الصناعي بها والعلاقة بينها وبين التنبؤ علاقة طردية فكلما تضخمت البيانات سهل التوقع[2]، على الرغم من أن الأمر يبدو آمنا ومحافظا على سلامة الأشخاص والعدالة بادئ الأمر إلا أن نتائجه غير منزهة عن الخطأ وهذا ما يدل عليه تقرير راند rand.[3] الذي نشرته والذي فصلت الكلام فيه حول تقنية التنبؤ الاجرامي.

فكانت المخاوف المتمحورة حول هذا النظام احتمالية إدانة الخوارزميات ومتابعتها لأشخاص على وقائع لم يقترفوها بعد، واحتمالية أن يتم استهداف الأفراد الملونين ظلما مما يطعن في ما يسمى بالعدالة الخوارزمية الهادفة إلى معرفة الفعل الاجرامي والتنبؤ به قبل ارتكابه، هذا ما يجعلها تحمل في طياتها العديد من المخاطر منها عدم الدقة في التنبؤ بالنتائج وهذا نتيجة ادخال بيانات غير صحيحة بسبب التحيز البشري

[1] عبد القادر غونيول، "السلوك البشري يفرض عنصريته على الذكاء الاصطناعي"، https://www.aa.com.tr/، نشر بتاريخ 2020/06/28، تاريخ الاطلاع 2022/08/23، 15:43.

[22] محمود سلامة عبد المنعم الشريف، "الطبيعة القانونية للتنبؤ بالجريمة بواسطة الذكاء الاصطناعي ومشروعيته"، المجلة العربية لعلوم الأدلة الجنائية والطب الشرعي، العدد 3، 2021، جامعة نايف العربية للعلوم الأمنية، السعودية، ص343.

[3] محمود سلامة عبد المنعم الشريف، مرجع سابق، ص 349.

مما يؤدي إلى حدوث تحيز في اختيار البيانات المكونة للخوارزمية نفسها لأن التحيزات الحسابية ما هي الا مرآة مكبرة للتحيزات البشرية.[1]

لهذا ناقش العديد من علماء القانون مسألة اعتماد النظم الخوارزمية التنبؤية وهذا ما تناولته تقارير مؤسسة راند (rand) التي ناقشت حدود استخدام الخوازميات التنبئية في انفاذ القانون المدني الأمريكي، كذا ناقش التقرير الصادر عن منظمة (propublica) اعتماد خوارزمية (compas) لتقدير العود للإجرام وأظهرت النتائج أنّ النظام كان يعطي نتائج منحازة على نحو منهجي مما أدى إلى تفاوت كبير في نتائج الأحكام الجنائية في المحاكم التي تستخدم هذه التقنية.[2]

ومؤخرا كشفت مجموعة من علماء جامعة شيكاغو عن خوارزمية قادرة على التنبؤ بالجرائم بدقة تبلغ 90 % فتحدد المواقع التي بها احتمالية وقوع الجرائم بنسبة أكبر، إلا أن اعتماد الذكاء الاصطناعي في التنبؤ بالجريمة والمجرم يدق ناقوس العدالة، إذ ذكر فينسينت ساثرلاند المدير المشارك لمركز العرق وعدم المساواة والقانون في جامعة نيويورك أنّ جودة الخوازميات تتعلق بجودة البيانات التي تغذى بها، وهذا يشكل مشكلة بالنسبة لسكان الولايات المتحدة الأمريكية إذ أن بيانات الشرطة هناك متحيزة تاريخيا فرجال الشرطة غالبا ما يقبضون ويتهمون من يقطن بالأحياء الفقيرة التي غالب سكانها ملونين وهذا ليس صحيح دائما، وهذا يعني أن معظم البيانات تبالغ في تمثيل الملونين والأحياء منخفضة الدخل، يقول في هذا ساثرلاند: "لديك بيانات مصابة أو مشوبة ببعض التحيز وهذا التحيز سيظهر على الجانب الآخر من التحليل لأن الخوارزمية تخرج لك ما تضعه فيها" لهذا السبب أوقعت برامج التنبؤ بالجريمة في حدوث مشكلات كبيرة ففي عام 2019 تم ايقاف العمل ببرنامج laser الذي يستخدم البيانات التاريخية للتنبؤ بالنقاط النشطة للجريمة وبرنامج بالانتير (PALANTIR) بعد أن أخضع أشخاص من ذوي أصول افريقية ولاتينية بغير عدل لمراقبة أشد.[3]

إذا لم يعد من الآمن الاعتماد الكلي على أنظمة التنبؤ الخوارزمي لمعرفة مواطن ارتكاب الجرائم، أو لمتابعة بعض الأفراد ممن اشتبه فيهم النظام كونه غير عادل، فحقيقة أن الخوارزميات تتعلم آليا من قاعدة البيانات المتوفرة والمحتمل تشبعها المسبق بالتحيزات البشرية، لا يجعل أي شك في عدم مصداقية النتائج

[1] عمر عبد المجيد مصبح، "توظيف خوازميات العدالة التنبؤية في نظام العدالة الجنائية الآفاق والتحديات"، المجلة الدولية للقانون، المجلد 10، العدد 1، 2021، ص 255.

[2] أوسوندي أ. أوسوبا، ويليام ويلسر الرابع، "مخاطر الذكاء الاصطناعي على الأمن ومستقبل العمل"، مؤسسة rand (www.rand.org) ، 2017.

[3] برانشو فيرما، "الذكاء الاصطناعي ومخاطر التحيز"، accronline.com 28 يوليو 2022، تاريخ الاطلاع 2022/08/23، 14:38.

المتوصل إليها، وتميزها بالعنصرية خاصة ضد الأشخاص ذوي البشرة السوداء، وضد قاطني المناطق الفقيرة النائية.

إن المتتبع للتقارير التي تضعها المؤسسات العالمية حول مسألة أخلقة الذكاء الاصطناعي والعيوب الناتجة عنها ومساسها بحقوق الانسان وحرياته الأساسية، والشكاوي التي ينشرها المستخدمون التي تلخص تجاربهم الفاشلة مع النظام يمكنه استقصاء نماذج أخرى تذكر على سبيل التمثيل لا الحصر:

3) أنظمة محركات البحث

يعتبر غوغل أحد أكبر محركات البحث على الأنترنيت، مما يعني أنه يتم الاعتماد عليه بنسبة معتبرة جدا، إلا أن النتائج التي يقدمها قد لا تكون مرضية دائما وبالأحرى قد تكون عنصرية بدرجة كبيرة، فمثلا اكتشف المستخدمون أنهم إذا ما كتبوا كلمة غوريلا في البحث سيكون المواطنون السود ضمن النتائج.[1]

وعند البحث عن كلمة طفل فإن النتائج ستكون أغلبها عائدة إلى أطفال بيض، إلا أن مسؤولي غوغل قد برروا موقفهم بأن نتائج بحث غوغل تعكس المحتوى الأكثر شيوعا على الانترنيت،[2] إن هذا التصريح الذي قدمته الشركة وإن كان يدعمه التعلم الآلي للخوارزميات من البيانات المتوافرة مسبقا والتي تعكس توجه الأفراد وآرائهم، إلا أنه يستوجب اصلاحا تقنيا للمشكلة.

4) أنظمة المر افق الخدماتية

مثلا نجد في صنابير المياه الأوتوماتيكية في المطاعم وغيرها عادة ما يشتكي السود من عدم اطلاق تلك الصنابير للمياه والصابون والسبب أن تلك الأجهزة لم تصمم بشكل يجعل متحسساتها تستجيب للون البشرة الداكنة.[3] إن هذه المعاملة العنصرية في أبسط المرافق كالفنادق يدل على تجذر مشكلة التحيز الرقمي العنصري في أنظمة أكبر وفي مرافق ومجالات أكبر.

---

[1] طه الراوي، "عنصرية التكنولوجيا...عن أسباب وأشكال التحيز الرقمي"، مرجع سابق.

[2] كونيل جاي، "الذكاء الاصطناعي: خوازميات العنصرية"، ترجمة: أحمد الشاعر، alaraby.co.uk، نشر بتاريخ 12 يوليو، 2017، تاريخ الاطلاع 2022/08/24، 13:38.

[3] طه الراوي، "عنصرية التكنولوجيا...عن أسباب وأشكال التحيز الرقمي"، مرجع سابق.

5) أنظمة برامج التواصل الاجتماعي

إن وسائل التواصل الاجتماعي ورغم تعديلها لخوارزمياتها للحد من مقاطع الكراهية إلا أنها تبقى بيئة خصبة لنمو الخطاب المحرض على العنصرية والكراهية.[1]

نستنتج مما سبق أن العنصرية الرقمية لا تقتصر على مجرد تمييز النساء عن الرجال، أو البيض عن السود في التوظيف مثلا، إن المسألة أعمق بكثير، لقد أصبحت المشكلة تمس كل جوانب الحياة بدأ من التمتع بالمرافق الخدماتية، إلى أحدث تقنية قد يصل اليها أي برنامج.

3. أسباب التحيز الرقمي

كشفت دراسة أطلقتها جامعة جونز هوبكنز ومعهد جورجيا للتكنولوجيا وباحثين من جامعة واشنطن بأن أنظمة الذكاء الاصطناعي المعتمدة على التعلم الآلي من شأنها ابداء نتائج متحيزة على أساس الجنس والعرق، يقول في هذا أندرو هوندت "نحن معرضون لاختراع جيل من الروبوتات العنصرية والمتحيزة جنسيا"[2] بالرغم من أن السبب الرئيس للتحيز في خوازميات أنظمة الذكاء الاصطناعي هو التحيز الموجود عند المستخدمين بالدرجة الأولى، إلا أنه توجد طرق أخرى من خلالها تكتسب الخوازميات منها عنصريتها:

1.3 البيانات المتحيزة

إن الخوازميات من ناحية المنطق لا تمتلك أي طابع تحيزي تلقائي فالتحيز ما هو إلا صفة بشرية وبما أن الذكاء الاصطناعي يحاكي الذكاء البشري عن طريق التعلم الآلي من البيانات القاعدية المغذية له جعلت هذه الأنظمة ترث ذلك التمييز العنصري، [3] وحقيقة أن الانسان من شأنه أن يُورّث الآلة عنصريته رغم علمه وإدراكه التام بما كلفته عنصريته على مر العصور، لهو أمر خطير وكارثة كبيرة ودليل على أن هناك خللا بنيويا خطيرا في فكر الأمن الإنساني المستقبلي.[4]

وبما أن الخوازميات أنشأها أشخاص لديهم مجموعاتهم الخاصة من التحيزات فإن الأفكار المسبقة من شأنها التأثير على كيفية عمل الذكاء الاصطناعي مما يؤدي قطعا إلى أنظمة لا تعمل الا لجزء من

---

[1] المرجع نفسه.

[2] "دراسة تكشف أن أنظمة الذكاء الاصطناعي تتخذ قرارات عنصرية ومنحازة جنسيا"، mc-doualiya.com نشرت في 2022/06/27، 11:46، تاريخ الاطلاع 2022/08/23.

[3] محمد فتحي، "الذكاء الاصطناعي ينافس الانسان في الانحياز"، www.alkhaleej.ae نشر بتاريخ 25 يوليو 2018، تاريخ الإطلاع 2022/08/22، 14:37.

[4] المختار السالم احمد السالم، في ظلال الحروف، الناشر شركة بريطانية e_kutub ltd ، ط1، لندن، ماي، 2019، ص 9 .

المستخدمين يشهون صانعي هذه الأنظمة،[1] إذا فتغذية نماذج التعلم الآلي بالبيانات التي تجمع من العالم الحقيقي من شأنها أن تجعل أكثر الأنظمة تتعلم أو تضخم التحيزات الموجودة مسبقا في هذه البيانات والتي يمكن أن تحتوي على تحيز قائم على أساس العرق أو الجنس أو الدين أو أي خصائص أخرى،[2]مثلا إذا ما تم البحث في محرك البحث (غوغل) للصور عن كلمة لاعب فإن اغلب النتائج ستظهر صورا متعلقة بكرة القدم في حين قد لا نجد صورة واحدة للاعب شطرنج، وهذا مرتبط بنوعية البيانات التي تعلم منها والتي غالبا ما تربط كلمة لاعب بلعبة كرة القدم.[3] وهذا ما نجده أيضا جليا مثلا في أنظمة التعرف على الوجه التي تم تدريبها على صور ذوي البشرة البيضاء بنسبة 90 مئة مقابل 10 بالمئة لذوي البشرة الداكنة.[4] ففي سنة 2016 اهتمت الباحثة جوي بيولامويني وهي باحثة أمريكية من أصل افريقي في معهد ماساتشوستس للتكنولوجيا بخوارزميات التعرف على الوجه اذ اكتشفت ان النظام لا يتعرف على أصحاب البشرة السوداء فأظهرت أبحاثها أن الخوازميات تم تدريبها على قواعد بيانات تحتوي معظمها على صور رجال بيض لذلك كان أدائها سيئا عندما يتعلق الأمر بصور النساء السوداوات أو الرجال الآسيويين مقارنة بالبيض.[5]

إذ يلجأ المبرمجون عادة في بناء هذه الأنظمة المتعلقة بالتعرف على الوجه على بيانات ضخمة متاحة على الانترنيت، غير أنّ هذه البيانات غير دقيقة ومنحازة بشكل علني مما يعني أن اعتمادها في بناء الخوازميات من شأنه أن يشبع بنفس التحيزات، وللتأكد من هذا اعتمد الباحثون على شبكة عصبية تسمى **clip** وهو نظام يقوم بمطابقة الصور بناء على البيانات الموجودة في النت، وطلب منه وضع صور عشوائية في صندوق بني يحمل في كل مرة صفة معينة مثلا وضع الأشخاص في الصندوق البني، وضع الطبيب في الصندوق البني، وضع المجرم في الصندوق البني، ووضع عمال النظافة في الصندوق البني، ومن أبرز النتائج الملفتة للانتباه أن النظام اختار الذكور اكثر بنسبة 8% ، واختير الرجال البيض والآسيويون أكثر من غيرهم، واختيرت النساء ذوات البشرة السوداء أقل من باقي الفئات، وحدّد الرجال المجرمون على أنهم سود بنسبة 10 % أكبر من الرجال البيض، وتم تحديد الرجال اللاتينيين على أنهم عمال نظافة بنسبة 10 بالمية اكثر من

---

[1] دان دايلي، "للتغلب أولا على التحيز في الذكاء الاصطناعي يجب أولا الاعتراف بوجوده"، fortunearabia.com نشر بتاريخ 10 فيبراير 2022، تاريخ الاطلاع 2022/08/22، 10:10.

[2] "كيف يمكن التغلب على تحيز الذكاء الاصطناعي"، مرجع سابق.

[3] محمد فتحي، مرجع سابق.

[4] طه الراوي، "عنصرية التكنولوجيا عن أسباب وأشكال التحيز الرقمي"، www.noonpost.com نشر بتاريخ 2020/06/12، تاريخ الاطلاع 2022/08/22، 06:44.

[5] ar.unesco.org، 2022/08/22، 8:57.

الرجال البيض، واختار النظام كطبيب الرجال بنسبة أكبر من جميع النساء في جميع الأعراق.[1] كذلك تنسب برامج الترجمة جنسا معينا إلى المهن على سبيل المثال ترجمة كلمة ذ دكتور (غير مجنسة) إلى الدكتور (مذكر) بالفرنسية، أو ذ نووس إلى ممرضة بالفرنسية مؤنثة.[2]

وكمثال آخر عن تأثير البيانات المتحيزة في عنصرية الخوازميات أنه في أحد المدن تم اعتماد خوارزمية كان الهدف منها هو الكشف عن الأخاديد والحفر الموجودة في الشارع بناء على معطيات وبيانات يتم جمعها من مستخدمي الهواتف الذكية، لكن لم يتم رصد جميعها بسبب التوزيع الديموغرافي آنذاك لمستخدمي الهواتف الذكية فاسفر عن ذلك نقص الخدمات لبعض المجتمعات المحرومة أي أدى إلى حرمان المواطنين الأقل ثراءً من الحصول على خدمات الصيانة في تلك المدينة.[3]

وصرحت العالمة أود بارنهايم لليونسكو في مقابلتها معها: "لم نفاجأ كثيرا عند عثورنا على ميل للتمييز الجنسي في التكنولوجيات بصفته موجودا في العديد من المجالات الأخرى وبالتالي ليس هناك من مبرر لعدم وجوده في مجال الذ الاصطناعي" إلا أن النتائج السلبية قد تكون أكثر حدة من المعتاد فنجد أن القوالب النمطية التي تحملها الخوارزميات يمكن أن تقوم باستبعاد النساء عند الترشيحات مسبقا في المناصب التقنية مثلا،[4] هناك العديد من الأمثلة التطبيقية الأخرى التي تفند حقيقة أن التعلم الممكنن من البيانات المتحيزة يؤدي إلى تحيز الخوازميات، فهذه الأخيرة تعكس بالضرورة جودة البيانات التي تغذى بها وبما أنها مستمدة من البشر فغالبا ما ستتعلم التحيزات منهم سواء كانت جيدة أم سيئة.[5]

وعليه يعتبر السبب الأساسي وراء تحيز نتائج أنظمة الذكاء الاصطناعي هو البيانات المتحيزة التي تغذي البرنامج والتي تتوافر بوفرة على الانترنيت، هذه الأخيرة تشكل التوجهات التي يتبناها الأفراد وأفكارهم وتوجهاتهم مما يعني عدم حياديتها وعدم نقاهتها من العنصرية ضد بعض الأجناس أو الأعراق.

---

[1] "دراسة تكشف أن أنظمة الذكاء الاصطناعي تتخذ قرارات عنصرية ومنحازة جنسيا"، مرجع سابق.

[2] ar.unesco.org، مرجع سابق.

[3] أوشونديه أوشوبا، ويليام ويلسر الرابع،"ذكاء اصطناعي بملامح بشرية مخاطر التحيز والاخطاء في الذكاء الاصطناعي"، مؤسسة rand، (www.rand.org، كاليفورنيا،2017، ص4.

[4] ar.unesco.org، مرجع سابق.

[5] " أوشونديه أوشوبا، ويليام ويلسر الرابع، مرجع سابق، ص 17.

3.2 التعميمات

إن تعميم المخاطر وبثها داخل الخوازميات سبب آخر للعنصرية فمثلا مسألة أن معظم ذوي البشرة السوداء خطيرين في هذه الحالة سوف تطلق الخوازميات تحذيرات خاطئة ضد أناس أبرياء،[1] إن هذه التعميمات تشكل خطرا كبيرا على حرية الأفراد وأمنهم إذ من شأنها أن تجعل ذوي البشرة السوداء دوما محل شك واتهام، فتتعزز العنصرية البشرية ضدهم بالعنصرية التقنية.

3.3 الخوازميات المعيبة

إن الاعتماد الكبير على خوازميات اتخاذ القرار قد يشكل خطرا على خصوصية وأمن الأفراد وقد سبق أن حذر مكتب البيت الأبيض لسياسة العلوم والتكنولوجيا من امكانية ترميز التمييز في القرارات الممكننة، يقول ديديو في هذا "قد تكون الخوازميات مثالية من ناحية الرياضيات ولكن تثير الاشكالية من الناحية الأخلاقية".[2]لذا فإن من مصادر التحيز في أنظمة الذكاء الاصطناعي الخوازميات نفسها إذ يمكن للخوارزميات المعيبة تضخيم التحيزات من خلال التغذية الرجعية، مثال ذلك خوارزمية تضمين الكلمات إذ ترى الأسماء الأوروبية جيدة وأسماء الأفارقة الأمريكيين بغيضة.[3] مثال ذلك تطبيق google translate إذ يعتمد على الترجمة بالضمائر المذكرة وليس المؤنثة، ويرجع السبب في هذا إلى أن نسبة الضمائر المذكرة في مصادر نصوص اللغة الانجليزية يبلغ ضعف الضمائر المؤنثة فيه، وفي كل مرة يرتكب التطبيق هذا الخطأ في الترجمة يزداد التكرار النسبي للضمائر المذكرة على شبكة الانترنيت.[4]

4.3 عدم التنوع

صرحت الباحثة فلورا فانسان بدورها إلى اليونسكو أثناء المقابلة التي أجرتها معها هي وزميلتها أود برنهايم أن الفرق العلمية تفتقر إلى التنوع، إذ أثبتت مؤرخة العلوم الأمريكية لوندا شيبنجر أنه كلما ازداد عدد النساء في الفريق ازداد احتمال مراعاة مسألة النوع في الدراسة نفسها، والأمثلة على التمييز في مجال

---

[1] طه الراوي، "عنصرية التكنولوجيا عن أسباب وأشكال التحيز الرقمي"، مرجع سابق.

[2] أوشونديه أوشوبا، ويليام ويلسر الرابع، مرجع سابق، ص2.

[3] بايرام كاناريا وآخرون، الدين والقيم والأخلاق، اسطنبول، ص 166.

[4] "تصميم الذكاء الاصطناعي بحيث يكون منصفا"، m.annabaa.org، تاريخ النشر 2021/12/21، تاريخ الاطلاع 2022/08/24، 17:57.

البحث عديدة منها في مجال صنع الأدوية إذ أنه يتم تجريب الأدوية على الفئران الذكرية لنقص الهرمونات بها مما يسهل من تقليل الآثار الجانبية.[1]

وهذا ما ذهب إليه العديد من الباحثين على رأسهم جوي بولامويني إذ أقرت بأن أحد أسباب عنصرية الخوازميات هو عدم تنوع الموظفين التقنيين في الشركات التكنولوجية بشكل كاف، فحسب التقرير الذي نشرته غوغل في 2016 فإن موظفيها التقنيين 19 % منهم نساء و1 %بالمئة فقط من ذوي البشرة السوداء، ونشرت مايكروسوفا أن 17.5 % من موظفيها التقنيين نساء و 2.7 % فقط من أصحاب البشرة السوداء، كذلك أظهرت التقارير أن نسبة الموظفات التقنيات في شركة فيسبوك 17% والسود 1 % فقط.[2]

هناك عوامل تقنية أخرى تعزز من تحيز الخوازميات مثل تفاوت حجم العينة إذ أن خوازميات التعلم الآلي هي في الأصل وسائل احصائية فهي هنا تعاني مشاكل في معالجة التفاوتات في حجم العينة، وقد تواجه الخوارزمية أيضا صعوبة في تسجيل تأثيرات ثقافية معينة عندما يكون السكان مقسمين إلى شرائح متعددة.[3]

إن عملية اتخاذ القرار من قبل أنظمة الذكاء الصناعي يتطلب تغذية الخوازميات بالبيانات القاعدية المتوافرة على مستوى الانترنيت، مما يعني أن أهم عنصرين فاعلين في العملية التي من شأنها أن تنتج لنا قرار متحيزا هما: البيانات والخوارزميات فكل ما من شأنه التأثير على حيادهما يعد سببا لتحيز الذكاء الصناعي، وهذا بدأً من المستخدمين الذين تتجسد عنصريتهم وارائهم التمييزية على شكل بيانات مما يجعل هذه الأخيرة غير حيادية هي الأخرى، إلى المبرمجين أنفسهم أين يفتقر الوسط البحثي إلى التنوع بينهم، وإلى افتقارهم شخصيا للوعي بضرورة التنوع الحاصل في الواقع وتوظيفه في الرقمنة.إلى الخوارزميات التي قد تكون معيبة وغيرها من الأسباب الأخرى.

4. تدابيروآليات الحد من تحيز أنظمة الذكاء الاصطناعي.

تقول الباحثة أود بارنهايم "إن اشكالية التمييز ظهرت في هذا الحقل في وقت مبكر للغاية إذ لم تكد تمر ثلاث سنوات على بداية الحمى الخوارزمية حتى بادر المنبهون إلى المخاطر بجلب الانتباه نحو المعاملة

[1] اليونسكو، مرجع سابق.

[2] كونيل جاي، الذكاء الاصطناعي خوازميات العنصرية، ترجمة أحمد الشاعر، alaraby.co.uk ، تاريخ النشر 12 يونيو 2017، تاريخ الاطلاع 2022/08/24، 13:38.

[3] أوشونديه أوشوبا ووليام ويلسر الرابع، مرجع سابق، ص 19.

التفاضلية لبعض الخوازميات وهي مسألة أصبحت منذ وقتها موضوعا في حد ذاته داخل المجتمع العلمي يغدي عديد النقاشات وتترتب عنه أعمال بحثية عن التحيزات وتأثيرات الخوازميات من وجهة نظر أخلاقية وحسابية واعلامية، ولقد أصبحنا في الآونة الأخيرة نجد أثرا لهذا الوعي في الصحافة العامة، وحتى لئن لم يتم حل جميع المشكلات فهي على الأقل أصبحت معروفة وهو ما قد يمهد الطريق بمعالجتها".[1]

إذا فالانتشار الاعلامي الواسع لهذه المشكلة أصبح يؤكد _بشكل لا مجال للشك فيه_ أن التحيز الرقمي ليس مجرد خلل تقني أو حاصل تحيز بشري سابق، بل هو مسألة خطيرة تستدعي تدخل المسؤولين والباحثين بصفة عاجلة لتأثيرها الكبير على حقوق الانسان وحرياته الأساسية لإيجاد حلول قاطعة جذرية، أو على الأقل التخفيف من حدة الآثار وتقليلها، ويمكن القول أن الآليات والتدابير التي يجب اتخاذها يمكن اجمالها في ما يلي:

1.4 وضع اجراءات تنظيمية

صرحت منظمة اليونسكو أن هناك جانبا اخرا للأنظمة الذكاء الاصطناعي يتمثل في تعمق الفجوات وعدم المساواة لهذا كان لزاما وضع أطر تنظيمية دولية ووطنية بغية ضمان عودة هذه التكنولوجيات بالنفع لا الضرر على البشرية، لذلك تقترح اليونسكو وثيقة تقنينية عالمية شاملة بغية تأسيس الذكاء الاصطناعي على قاعدة أخلاقية تحمي وتعزز حقوق الانسان فتكون بذلك بوصلة أخلاقية تسمح بإرساء احترام سيادة القانون في العالم الرقمي،[2] في نفس الاطار صرحت مؤسسة راند بالآتي "ستستمر مخاطر الخطأ والتحيز في الخوارزميات والذكاء الاصطناعي طالما يؤدي وكلاء الذكاء الاصطناعي أدوارا بارزة أكثر من أي وقت مضى في حياتنا دون تنظيمهم"[3] و "ينبغي وجود جرعة صحية من القيود التنظيمية من أجل تحقيق التوازن مع التوجه نحو تدابير معالجة التحيز في الخوازميات"[4] لهذا قامت العديد من المنظمات الدولية قد عقدت مؤتمرات تناقش الأمر، والعديد من الشركات والمؤسسات وضعت مدونات خاصة بأخلاقيات الذكاء الصناعي وقامت العديد من الجهات بوضع التقارير حول الأخلاقيات مثل الاتحاد الأوروبي، ومعهد المستقبل من أجل الحياة future for life institute، ومعهد مهندسي الكهرباء والالكترونيات institue of electrical

[1] اليونسكو، مرجع سابق.

[2] عن اليونسكو.

[3] أوشونديه أوشوبا ووليليام ويلسر الرابع، مرجع سابق، ص 25.

[4] المرجع نفسه، ص 24.

and electronics engineers، واقترحت المجموعة الأوروبية المعنية بأخلاقيات العلوم والتكنولوجيا الجديدة اطارا أخلاقيا يضم جملة من القيم والمبادئ المعرضة للخطر بسبب الذكاء الاصطناعي.[1]

وقد أصدرت اليونسكو في 23 نوفمبر 2021 توصية خاصة بأخلاقيات الذكاء الاصطناعي ومما جاء فيها:

- _ إن حقيقة وجوب التصدي لأجل أخلقة أنظمة الذكاء الاصطناعي لا يجب أن يكون حاجزا يعيق الابتكار والتطور بل لابد أن يكون حافزا على البحث بطريقة أخلاقية تربط تكنولوجيا الذكاء الاصطناعي بحقوق الانسان وحرياته الأساسية وبالقيم والمبادي الأخلاقية.
- _ وجاء فيها "تؤدي خوازميات الذكاء الاصطناعي إلى ايجاد تحديات أخلاقية جديدة إذ يمكن أن تتسبب هذه الخوازميات في اعادة انتاج اوجه التحيز الموجودة وتعزيزها وقد يتسبب هذا الأمر بالتالي في تفاقم أشكال التمييز والأفكار أو الأحكام المسبقة والقوالب أو الصور النمطية الموجودة أصلا".
- _ من القيم الواجبة التطبيق التي نصت عليها التوصية هي احترام حقوق الانسان والحريات الأساسية وكرامة الانسان لذلك لا بد من حمايتها طوال فترة نظام الذكاء الصناعي وحمايتها وفقا للأحكام القانون الدولي منها أحكام القانون الدولي لحقوق الانسان.
- _ يجب أن توفر الوسائل التكنولوجية أدوات جديدة داعمة لحقوق الانسان والدفاع عنها ويجب ألا تؤدي إلى انتهاك هذه الحقوق.
- _ وينبغي للجهات الفاعلة المعنية بالذكاء الاصطناعي أن تبذل كل الجهود المعقولة للحد قدر المستطاع من التطبيقات والنتائج التمييزية أو المتحيزة وتفادي تعزيزها أو تخليدها....ويجب أن يكون هناك سبيل فعال لطلب الانصاف عند التعرض للتمييز والقرارات الخوارزمية المتحيزة.
- _ "ينبغي على الدول الأعضاء وضع استراتيجيات لإدارة البيانات تضمن تقييم نوعية البيانات المستخدمة لتدريب نظم الذكاء الاصطناعي".
- _ دعت الى ضرورة "اشتمال السياسات الرقمية الوطنية على خطة عمل للمساواة بين الجنسين ووضع سياسات ملائمة لدعم الفتيات والنساء بشأن مسائل تضم مثلا الأعداد للالتحاق بسوق العمل من أجل ضمان عدم استبعادهن من الاقتصاد الرقمي القائم على الذكاء الص".

[1] "تطوير استراتيجية للذكاء الاصطناعي دليل وطني"، الأمم المتحدة، الاسكوا، www.unescwa.org، ص 16.

- _دعت إلى أنه "ينبغي أن تضمن استخدام هذه الوسائل بطريقة تتيح سد الفجوات وازالة الفوارق الموجودة بين الجنسين وتضم هذه الفجوات والفوارق الفارق في الأجور بين الجنسين وعدم المساواة في تمثيل الجنسين في مهن وانشطة معينة".
- _ "ينبغي للدول الأعضاء أن تضمن تفادي استنساخ القوالب أو الصور النمطية الجنسانية وأوجه التحيز التمييزية في نظم الذكاء الاصطناعي بل ينبغي لها أن تسعى إلى الوقوف على هذه الأمور وأن تبادر إلى التصدي لها"[1]

2.4 تحديد التحيزات المحتملة في المرحلة الأولى من بناء النظام:

بعد تزايد الوعي بوجود التفاوت الاقتصادي في المجتمع والعرق والجنس وغيرها كان لابد على الشركات من التأكد من أن منتجات الذكاء الاصطناعي التي بنونها خالية من التحيز قدر الإمكان والخطوة الأساسية لذلك هي تحديد التحيزات المحتملة في المراحل الأولى من تطوير المنتج، تقول في هذا آليس شيانغ رئيسة مكتب أخلاقيات الذكاء الاصطناعي في مجموعة سوني في مؤتمر العصف الذهني الخاص بالذكاء الاصطناعي (brainstorm al) الذي نظمته fortune مؤخرا في بوسطن "لا توجد خوارزمية غير متحيزة نهائيا، لذلك أقول دائما لوحدات الأعمال أن طلب اجراء تقييمات النزاهة ليس لأننا نعتقد أنهم ارتكبوا خطأ ما أو أن منتجهم على وجه الخصوص يعاني من مشاكل ما بل لأننا نرغب دائما في التحقق من هذه الصفة"، [2] وبينت أن وحدات الأعمال لا يريد أحد منها انتاج منتجات منحازة ولكن من الصعب حقا معرفة التقنيات الصحيحة والمعايير الصحيحة للتخفيف من التحيز الناتج عنها.[3]

وضحت كذلك الدكتورة هانية محموديان وهي خبيرة في أخلاقيات الذكاء الاصطناعي في شركة داتا روبوت (datarobot) أن الذكاء الاصطناعي لن يكون أبدا مثاليا إذ يتعلق الأمر بمراقبته بشكل أساسي في كل خطوة ضمن عملية بنائه للتأكد من أن البرنامج لا يكتسب صفة التحيز، هذه المراقبة تدفعنا للعلم هل البرنامج جاهز لوضعه في مرحلة الانتاج أم لا بد من التراجع وجمع بيانات أفضل حتى يصبح النظام جاهزا للنشر.[4]

[1] التوصية الخاصة بأخلاقيات الذكاء الاصطناعي، اليونسكو.

[2] 2022/08/22/10:24 موقع fortunearabia.comبقلم دان دايلي 10 فيبراير 2022

[3] دان دايلي، مرجع سابق.

[4] دان دايلي، مرجع سابق.

ولمعالجة القصور الملاحظ في أنظمة التعرف على الوجه دربت بعض المؤسسات خوارزمياتها حتى تصبح أكثر شمولا، كما طورت بعض الشركات أدوات تمكنها من تقييم الخوازميات ورصد التحيزات التي قد تشوبها، ودعى بعض الباحثين أمثال جوي إلى تطوير مدونة تقوم على نبذ التحيز وتعميمها، ومن المبادرات القائمة اليوم العمل الذي تقوم به مجموعة data for good التي تسعى لوضع الخوازميات لخدمة المصلحة العامة ولقد قامت ايضا بصياغة ميثاق اخلاقي يعرف ب "قسم أبرقاط لعلماء البيانات" والذي وضع جملة من المعايير التي يجب التثبت منها قبل تطبيق اي خوازمية لعدم وجود تمييز فيها.[1]

3.4 رقمنة القيم الأخلاقية

إن من الأهداف التي يسعى لها الباحثين هي ضمان موثوقية الخوازميات بأعلى درجة ممكنة وجعلها في نفس الوقت متماشية مع قيم المجتمع ومبادئه[2]، أي جعلها غير خطيرة، وقد صرح البروفيسور ستيوارت رسل بأن الحل النموذجي لتجنب خطر التكنولوجيا هي بتحويل القيم الإنسانية إلى كود برمجي وأضاف: "أصبحت الروبوتات اليوم تقوم بمهام معقدة لم نكن لنتخيل أن يقوم بها أحد غيرنا نحن البشر، الأمر الذي بات يقتضي منا التفكير جديا في ترجمة قيمنا وأخلاقنا إلى لغة ذكاء اصطناعي نستطيع تلقينها لتلك الآلات الذكية"،[3]وهذا يتطلب خبراء في القيم الإنسانية كالفلاسفة وعلماء النفس وغيرهم، إلا أن التحدي الأكبر يتمثل أولا في تحديد القيم الأخلاقية الواجبة التعليم ثم برمجتها[4]، إذ من المعروف أن قطاع البحث في الذكاء الاصطناعي تابع للقانون الخاص وعليه من الممكن أن يكون له أفكار عن الأخلاقيات لا تتماشى بالضرورة مع التي يفضلها المجتمع، وحتى يتم تفادي الفجوة التي من الممكن أن تحدث لا بد من الحرص على مراعاة الطابع الأخلاقي للذكاء الاصطناعي عبر مناقشة موضوعية تأخذ بعين الاعتبار وجهات النظر المتعلقة بالأخلاق التي قد يثيرها الذكاء الاصطناعي في المجتمع، [5] أما بالنسبة لنوعية القيم التي من الممكن تلقينها للذكاء الاصطناعي حتى يكتسب سلوكا اخلاقيا نجد أن الباحثان تي وي آنغ ودفنا فينهولز قد وضحا الأمر

[1] اليونسكو، مرجع سابق.

[2] مصطفى سيسي، "من أجل استخدام ديموقراطي للذكاء الاصطناعي في افريقيا" رسالة اليونسكو، اليونسكو، ص 20.

[3] ثائر سوقار، "ثورة روبوتات تهدد مستقبل البشر"، www.skynewsarabia.com، 2015/11/10، 21:26.

[4] المرجع نفسه.

[5] رسالة اليونسكو، مرجع سابق، ص 31.

وبينا أنه إن كانت تلك القيم والأخلاق لصيقة وملازمة بالجنس البشري فإنها غير قابلة للبرمجة أما إذا كانت للأخلاق صفة عقلانية فمن الممكن اخضاعها للبرمجة.[1]

اشار المحلل التقني في مقال له طه الراوي بعض الخطوات التي تقلل من عنصرية الرقمية:

4.4 التنوع:

إن التنوع المطلوب هنا نوعان: هو تنوع في فرق البحث والابتكار بحيث تشمل الباحثين الاكاديميين على اختلاف جنسهم وعرقهم ولونهم ودينهم، وادراك المصممين أنفسهم بضرورة التنوع الحاصل في المجتمع سواء على مستوى الثقافات او غيرها.

التنوع على مستوى فرق البحث:

ويقصد به وجود أشخاص متعددي الألوان في الشركات المطورة لأنظمة الذكاء الاصطناعي حتى يتعرف الفريق على ما اذا كان المنتج يتعرف على الوجوه كلها أم لا في حالة انظمة التعرف على الوجوه.[2] فإذا لم يكن هناك فرق مشاركة مختلطة فهذا يعني أنه لا يوجد بالأساس اي ارادة لترميز المساواة في هذه الحالة فإن اي محاولة لتعديل خطوط البرمجة لن تكون كافية لحل أزمة التحيز الخوارزمي. [3] وهذا التنوع نصت عليه اليونسكو في التوصية الخاصة بأخلاقيات الذكاء الاصطناعي نددت بوجوب تشجيع النساء عل المساهمة والمشاركة في جميع مراحل دورة حياة أي نظام من نظم الذ الص وبوجوب تعزيز التنوع الجنساني في الوسائط البحثية الخاصة بالذكاء الصناعي عن طريق تحفيز النساء للإقدام على هذا المجال.[4] ولا يشمل التنويع وجود العلماء المبرمجين فقط بل يشمل ايضا علماء الاجتماع والقانونيين وغيرهم ممن لهم خبرة في قضايا التحيز حتى يساهموا بشكل ما في عملية تطوير الذكاء الصناعي. [5]

وعي المبرمجين بالتنوع:

تتطلب مسألة التحيز والأخطاء النظامية في وكلاء الذ الص (الخوازميات) منهجا مختلفا من مصممي الخوازميات إذ غالبا ما يكون اطلاعهم على القضايا والسياسات العامة والمجتمع غير كاف، وعادة ما تكون

[1] المرجع نفسه، ص 30.

[2] طه الراوي، مرجع سابق.

[3] اليونيسكو، مرجع سابق.

[4] التوصية الخاصة بأخلاقيات الذكاء الصناعي، اليونسكو، ص 21.

[5] طه الراوي، مرجع سابق.

الخصائص الديموغرافية التي يدمجونها غير متنوعة وعليه فوعي المبرمجين بالتنوع سيساعد في التقليل من اثار المشكلة.[1]

ويرى الباحث المتخصص في الذكاء الاصطناعي ابراهيم المسلم أنه كي نصل إلى ذكاء صناعي منصف بحق جميع المستخدمين لابد على المبرمجين من مراعاة الفروق والاختلافات واعداد بيانات الآلة وتعديلها وفقا للثقافة والهوية ، وكذلك المستخدمين إذ أن "عنصرية الذ الص امر يمكن التغلب عليه عندما نتغلب كمستخدمين على عنصريتنا". [2]

5.4 فحص البيانات المستخدمة

يجب فحص البيانات المستخدمة لتدريب وتغذية الذكاء الاصطناعي بشكل نقدي من أجل تفادي التحيز[3] فحتى تكون الخوازميات أكثر ميلا إلى المساواة يجب التصرف ابتداء على مستوى قواعد البيانات حتى تكون شاملة وممثلة للسكان في تنوعهم وهذا ما عملت به بعض الشركات إذ أخذت بعين الاعتبار الاختلافات في الجنس والجنسية والمورفولوجيا. [4] ولابد لهذه البيانات التي دربت عليها أنظمة الذكاء الاصطناعي أن تتسم بالشمول.[5]

ولا بد أن تكون هذه البيانات ممثلة اذ يجب على المؤسسات فهم كيف تبدو البيانات الممثلة قبل جمعها للتدرب عليها وضمان عدم احتوائها على اقل قدر من التحيز. [6]

تمثل هذه البيانات توجهات الأفراد فلابد من أن نكون نحن منصفين "عنصرية الذكاء الاصطناعي وأخطاؤه المحرجة هي انعكاس لانحيازنا، فقد أصبح عنصريا مثل المجتمعات التي صنعته وأحيانا يتفوق على عنصرية صانعيه لأن أدوات التعلم الآلي تواصل الربط بين البيانات الأصلية للتغذية فتعلم العنصرية والتحيز اكثر فأكثر" مثل ما حدث مع تصريحات روبوت الدردشة الذي أطلقته شركة تويت رتاي التي كانت

[1] أوشونديه أوشوبا ووليليام ويلسر الرابع، مرجع سابق، ص 24.

[2] نجود سجدي، "هل يورث البشر عنصريتهم للذكاء الاصطناعي؟" independentarabia.com، نشر بتاريخ 31 يوليو، 2022، تاريخ الاطلاع 2022/08/22، 14:59.

[3] طه الراوي، مرجع سابق.

[4] اليونسكو، مرجع سابق.

[5]، كيف يمكن التغلب على تحيز الذكاء الاصطناعي تقنيات وأدوات"، مرجع سابق.

[6] المرجع نفسه.

تصريحاته عنصرية وداعية للكراهية وهي عبارة عن برنامج يجسد فتاة مراهقة يتراوح عمرها بين 18 و 24 يمكن الدردشة معها على موقع تويتر.[1]

لهذا السبب قام معهد ماساشوستس الأمريكي للتقنية أم آي تي، مستودعا الكترونيا يضم مجموعة من البيانات المكونة من 80 مليون صورة من على موقعه، وهذا بعدما تأكد أن مجموعة البيانات تعد أكبر حاضنة للتحيز والتمييز ضد بعض الأعراق والأقليات والأجناس داخل عالم الذكاء الاصطناعي وهذا حتى لا يتم استعمالها مجددا في تغذية أنظمة الذكاء الاصطناعي.[2]

وايضا عادة يلجأ المصممون الى حذف المتغيرات الحساسة من البيانات التدريبية كمحاولة منهم للتخلص من التحيز الناتج في النظام إلا أن اخفاء المتغيرات الحساسة غالبا لا يحل المشكلة وهذا ما وضحه باروكاس وسيلبت إذ أن وسائل التعلم الذاتي في الخوارزمية من شأنه استنتاج متغيرات خفية عن طريق الاحتمالات[3]

6.4 تقليل الاقصاء

وهي ميزة تقوم على تقليل عدد المتغيرات المدخلة التي لا تحتوي على تباين كاف للتأثير على النتائج الى نماذج الذ الص بهدف تحسين التنبؤ بالنتائج[4].

7.4 عدم الاعتماد على الخوازميات فقط

إذ لابد من ابقاء الأفراد القائمين على تطوير الذكاء الاصطناعي على اطلاع بكل ما يتعلق بالنظام ليتمكنوا من التعرف بفعالية على أنماط التحيز غير المقصود مما يخلف نموذج تعلم آلي حيادي ويمكن أن يساعد توثيق حالات التحيز عند حدوثها وتحديد كيفية العثور عليها على ضمان عدم تكرارها.

4.8 الوعي والشفافية

"من أجل مكافحة تحيز الخوازميات، سيكون من المفيد تثقيف الجمهور وتسليحه بالوعي بأن الخوازميات يمكن أن تؤدي غلى نتائج غير منصفة، وهذا يختلف عن مطالبة المستخدمين بفهم اليات العمل

[1] بايرام كاناريا، مرجع سابق، ص 165.

[2] جمال محمد غيطاس، "مستودع ال 80 مليون صورة، أشهر حتضنة للتحيز والتمييز بأنظمة الذكاء الاصطناعي"، WWW.JOSOR.ORG، 13:33/08/24، 2022.

[3] أوشونديه أوشوبا ووليام ويلسر الرابع، مرجع سابق، ص 21.

[4] طه الراوي، مرجع سابق.

الداخلية لكل الخوازميات فلن يكون ذلك مجديا فسيكون من المفيد غرس جرعة صحية من الشكوك المدروسة لدى المستخدمين بما يكفي لتقليل حجم تأثير تحيز المكننة" اي أن الوعي والشفافية مفيد جدا هنا.[1]

_ وفي ما يلي أدوات للتغلب على تحيز الذ الص:

- **FAIRLENS** ي حزمة تحدد التحيز بشكل سريع وهي أحد مكتبات بايثون والتي تستخدم في اكتشاف النحيز وقياس الانصاف في البيانات بشكل تلقائي.
- **AEQUITAS** وهي ادوات مرنة تدقق في تحيز الذكاء الاصطناعي طوره مركز علوم البيانات والسياسة العامة في جامعة شيكاغو بالولايات المتحدة.
- **TCAV** هو نظام أعلن عنه سوندار بيتشاي الرئيس التنفيذي لشركة قوقل للكشف عن التحيز في نماذج التعلم الآلي.[2]

[1] أوشونديه أوشوبا ووليام ويلسر الرابع، مرجع سابق، ص 23.

[2] "كيف يمكن التغلب على تحيز الذكاء الاصطناعي تقنيات وأدوات"، مرجع سابق.

5. خاتمة

لكل تكنولوجيا جانب خطير ومظلم يقابل ما تقدمه للبشرية من منافع، وهكذا الأمر مع أنظمة الذكاء الاصطناعي، إلا أنه متى ما تعلقت الأخطاء بحقوق الأفراد ومستهم في حرياتهم فإن الأمر يصبح أكثر خطورة، ومن هذه التجاوزات التي وقعت فيها هذه الأنظمة التحيز الرقمي الذي تتعدد أسبابه من جهة وتتعدد التقارير التي تسعى لوضع آليات وحلول للحد منه من جهة أخرى، على إثر هذا خلصت الدراسة إلى جملة من النتائج هي في مجملها كالآتي:

- ✓ عرف العلماء التحيز الرقمي على أنه تحيز في خوازميات أنظمة الذكاء الاصطناعي وتفضيله لمجموعات دون أخرى.
- ✓ أصبح التمييز الرقمي امتدادا للتمييز البشري وهذا لعنصرية نتائجه تجاه بعض الفئات فنجدها في أنظمة التعرف على الملامح تتعرف على ذوي البشرة البيضاء دون البشرة السوداء، وفي المجال الأمني نجدها تميل إلى إدانتهم كذلك، فكان أكبر ضحية هم أصحاب هذه البشرة.
- ✓ تتعدد الاسباب التي تدفع بالخوارزميات للتحيز منها البيانات القاعدية، الخوارزميات، عدم تنوع الوسط البحثي.
- ✓ إن البيانات التي تعالج بها الخوارزميات غالبا ما تكون منتشرة على الأنترنيت أي أنها مشبعة بالتحيزات البشرية، لهذا أصبح التحيز الرقمي امتداد للتحيز البشري.
- ✓ وضعت منظمة اليونسكو وثيقة دولية مرجعية للدول في مجال الذكاء الاصطناعي وأخلقته تحت مسمى "التوصية الخاصة بأخلاقيات الذكاء الاصطناعي".
- ✓ من الآليات المقترحة والمعتمدة للتقليل من التحيز الخوارزمي هو فحص البيانات المستخدمة، عدم الاعتماد على الخوازميات فقط ، والوعي والشفافية.

التوصيات:

- ✓ ضرورة توعية المجتمعات (المستخدمة لكل أو بعض أنظمة معينة من الذكاء الصناعي) بمخاطر التحيز والعنصرية التي يمارسونها في الحياة الواقعية وكيف أنها تؤثر على الحياة الرقمية بالسلب.
- ✓ ضرورة تضمين القوانين الداخلية لأي دولة تتبنى الذكاء الاصطناعي جملة من النصوص والقواعد المتضمنة عواقب ردعية لكل من يعمد إلى انتاج أنظمة تتسم بالعنصرية.

- ✓ ابرام اتفاقيات دولية خاصة فقط بالذكاء الاصطناعي والتطوير التكنولوجي توافق عليها كل الدول التي بها شركات انتاج أنظمة الذكاء الاصطناعي، هذه الاتفاقيات تكون ذات بنود صارمة في معاقبة الشركات الكبرى المنتجة لأنظمة الذكاء الصناعي بشكل يعزز من العنصرية الرقمية ضد أي فرد.

6. قائمة المصادروالمراجع

1) "تصميم الذكاء الاصطناعي بحيث يكون منصفا"، m.annabaa.org، تاريخ النشر 2021/12/21، تاريخ الاطلاع 2022/08/24، 17:57.

2) "دراسة تكشف أن أنظمة الذكاء الاصطناعي تتخذ قرارات عنصرية ومنحازة جنسيا"، mc-doualiya.com نشرت في 2022/06/27، 11:46، تاريخ الاطلاع 2022/08/23.

3) "كيف يمكن التغلب على تحيز الذكاء الاصطناعي"، https://masaar.net، 2022/08/21، 22: 15.

4) ar.unesco.org

5) أوسوندي أ. أوسوبا، ويليام ويلسر الرابع،"مخاطر الذكاء الاصطناعي على الأمن ومستقبل العمل"، مؤسسة rand (www.rand.org) ، 2017.

6) أوشونديه أوشوبا، ويليام ويلسر الرابع،"ذكاء اصطناعي بملامح بشرية مخاطر التحيز والاخطاء في الذكاء الاصطناعي"،

7) بايرام كاناريا وآخرون، الدين والقيم والأخلاق، اسطنبول.

8) برانشو فيرما، "الذكاء الاصطناعي ومخاطر التحيز"، accronline.com 28 يوليو 2022، تاريخ الاطلاع. 2022/08/23، 14:38.

9) تطوير استراتيجية للذكاء الاصطناعي دليل وطني"، الأمم المتحدة، الاسكوا، www.unescwa.org.

10) ثائر سوقار، "ثورة روبوتات تهدد مستقبل البشر"،www.skynewsarabia.com 2015/11/10، 21:26.

11) جمال محمد غيطاس، "مستودع ال 80 مليون صورة، أشهر حتضنة للتحيز والتمييز بأنظمة الذكاء الاصطناعي"، WWW.JOSOR.ORG 2022،13:33/08/24

12) دان دايلي، "للتغلب أولا على التحيز في الذكاء الاصطناعي يجب أولا الاعتراف بوجوده"، fortunearabia.com نشر بتاريخ 10 فيبراير 2022، تاريخ الاطلاع 2022/08/22، 10:10.

13) السيد بدر، تصميم وتحليل الخوازميات، المكتبة العربية، د ط.

14) طه الراوي، "عنصرية التكنولوجيا عن أسباب وأشكال التحيز الرقمي"، www.noonpost.com نشر بتاريخ 2020/06/12، تاريخ الاطلاع 2022/08/22، 06:44.

15) طه الراوي، "عنصرية التكنولوجيا...عن أسباب وأشكال التحيز الرقمي"، www.noonpost.com، نشر بتاريخ 2020/06/12، تاريخ الاطلاع 2022/08/21، 15:36.

16) عبد القادر غونيول، "السلوك البشري يفرض عنصريته على الذكاء الاصطناعي"، https://www.aa.com.tr/، نشر بتاريخ 2020/06/28، تاريخ الاطلاع 2022/08/23، 15:43.

17) عمر عبد المجيد مصبح، "توظيف خوازميات العدالة التنبؤية في نظام العدالة الجنائية الآفاق والتحديات"، المجلة الدولية للقانون، المجلد 10، العدد 1، 2021.

18) كونيل جاي، "الذكاء الاصطناعي: خوازميات العنصرية"، ترجمة: أحمد الشاعر، alaraby.co.uk، نشر بتاريخ 12 يوليو، 2017، تاريخ الاطلاع 2022/08/24، 13:38.

19) محمد فتحي، "الذكاء الاصطناعي ينافس الانسان في الانحياز"، www.alkhaleej.ae نشر بتاريخ 25 يوليو 2018، تاريخ الإطلاع 2022/08/22، 14:37.

20) محمد مازن المصطفى، دمج خوارزميات التعرف على الوجه مع سرعة الطباعة من أجل المصادقة على الأنترنيت، رسالة ماجيستير، قسم الدراسات العليا علوم الويب، الجامعة الافتراضية السورية، سوريا، 2018/2017.

21) محمود سلامة عبد المنعم الشريف، "الطبيعة القانونية للتنبؤ بالجريمة بواسطة الذكاء الاصطناعي ومشروعيته"، المجلة العربية لعلوم الأدلة الجنائية والطب الشرعي، العدد 3، 2021، جامعة نايف العربية للعلوم الأمنية، السعودية.

22) المختار السالم احمد السالم، في ظلال الحروف، الناشر شركة بريطانية e_kutub ltd ، ط1، لندن، ماي، 2019، ص 9 .

23) مصطفى السيد حسين، "عنصرية الذكاء الصناعي... لماذا يبدو أصدقاؤك بملامح أوروبية عند استخدامهم تطبيقات وفلاتر الذكاء الاصطناعي؟"، www.arabicpost.net، نشر بتاريخ 2020/07/08، تاريخ الاطلاع: 2022/08/22، 14:35.

24) مصطفى سيسي، "من أجل استخدام ديموقراطي للذكاء الاصطناعي في افريقيا" رسالة اليونسكو، اليونسكو.

25) مصطفى سيسي، "من أجل استخدام ديموقراطي للذكاء الاصطناعي في افريقيا" رسالة اليونسكو، اليونسكو.

26) معجم البيانات والذكاء الاصطناعي انجليزي عربي، الهيئة السعودية للبيانات والذ الاصطناعي SDAIA، السعودية، ط1، 2022م.

27) مؤسسة rand،(www.rand.org)، كاليفورنيا،2017.

28) نجود سجدي، "هل يورث البشر عنصريتهم للذكاء الاصطناعي؟" independentarabia.com، نشر بتاريخ 31 يوليو، 2022، تاريخ الاطلاع /2022/08/22، 14:59.