

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي أي دور؟ وأي علاقة؟ Artificial Intelligence Algorithms and Social Media Content Ethics Which Role? which Relationship

د. عائشة كريكت *

جامعة محمد الصديق بن يحي جيجل - الجزائر، aichakriket18@gmail.com

تاريخ الاستلام 2022/05/26 تاريخ القبول 2022/12/23

الملخص

تهدف الدراسة الى الكشف عن الدور الممكن لخوارزميات الذكاء الاصطناعي في أخلقة محتوى مواقع التواصل الاجتماعي بمختلف انواعه، سواء المرتبط بالأخبار الكاذبة، الإرهاب او الكراهية وغيرها من المحتويات غير الأخلاقية التي أصبحت تهدد وعي وسلوك المستخدم عبر هذه المواقع. ويتأتى ذلك من خلال البحث في مفاهيم الذكاء الاصطناعي والخوارزميات، ثم تحديد وتحليل اهم الوظائف التي يمكن ان تقوم بها خوارزميات مواقع التواصل الاجتماعي، لنصل في الأخير الى فهم الدور الذي قد تلعبه هذه الأخيرة في حماية المستخدمين من التأثيرات السلبية طويلة المدى للمضامين غير الأخلاقية عبر هذه المنصات (فيسبوك، تويتر، يوتيوب وغيرها)، وفي نفس الوقت العلاقة بين هذين المتغيرين.

وقد توصلت الدراسة الى ان الخوارزميات يمكن ان يكون لها دور فعال في أخلقة مضامين مواقع التواصل الاجتماعي، ويمكن ان تكون كمصفاة للمضامين المسيئة والضارة بحيث تحول دون وصولها للمستخدمين، وهذا من خلال بعض الخدمات والإجراءات مثل: تعليق الحسابات والحذف التلقائي للمنشورات المسيئة، تحليل البيانات الوصفية للصور والفيديوهات، تحديد الكلمات والعبارات غير الأخلاقية وغيرها.

الكلمات المفتاحية: خوارزميات، الذكاء الاصطناعي، أخلاقيات، المحتوى، مواقع التواصل الاجتماعي.

Abstract:

The study aims to reveal the possible role of artificial intelligence algorithms in making the content of social networking sites ethical of all kinds, whether related to false news, terrorism or hate and other immoral content that threatens the awareness and behavior of the user through these sites.

This comes by researching on concepts of artificial intelligence and algorithms, identifying and analyzing the most important functions that social networking

algorithms can do, in order to reach finally an understanding of the role that may play in protecting users from the negative effects of unethical content through these platforms (Facebook, Twitter, YouTube, etc.), and at the same time the relationship between these two variables.

The study found that algorithms can have an effective role in ethics of the content of social networking sites, and can act as a filter for abusive and harmful contents, and this is through some services and procedures such as: account suspension and automatic deletion of offensive posts, metadata analysis For photos and videos, identify unethical words and phrases, and more..

Keys Words: algorithms, artificial intelligence, ethics, content, social networking sites.

1. مقدمة

أحدثت الإنترنت نقلة نوعية في المجتمعات بحيث تغيرت على المستوى البنيوي والوظيفي مختلف النشاطات المعرفية والإنتاجية والخدمية، وذلك بفضل الاندماج بين مختلف التكنولوجيات الحديثة وعلى رأسها تكنولوجيا الذكاء الاصطناعي القائمة على استخدام كميات هائلة من البيانات وتحليلها، والتي بدورها ساهمت في خفض تكاليف معالجة المعلومات، وتساعد أيضا الخوارزميات الحديثة القائمة على التعلم الآلي والبيانات الضخمة في القيام بعمليات عدة وتحليل بيانات هائلة في وقت أقل.

وتعد خوارزميات مواقع التواصل الاجتماعي نوعا هاما من عالم الخوارزميات عامة، بحيث تقوم بعدة وظائف مرتبطة بنشاط المستخدمين عبر هذه المواقع، والتي تندرج ضمن إدارة المحتوى وتفضيلات المستخدمين، ومن جهة أخرى فإن هذه الخوارزميات قد يكون لها دور مرتبط بالمحتوى غير الأخلاقي على مواقع التواصل الاجتماعي الذي تتنوع أشكاله، حيث ازدادت مؤخرا المضامين المسيئة للأفراد أو الجماعات المتعلقة أكثر بالكراهية، العنف، الإرهاب والتضليل، خاصة وان هذه المنصات تعد فضاءات خصبة لحرية التعبير وأحيانا دون مراعاة حقوق وحرية الآخرين ومعتقداتهم وانتماءاتهم.

وعليه انطلاقا من ذلك فإنه يمكن ان تقوم الخوارزميات نفسها بإدارة هذا النوع من المضامين، حيث تزيل ما هو غير أخلاقي على موقع فيسبوك او تويتر او يوتيوب، باعتبارها أكثر مواقع التواصل استخداما، فيمكن ان تحدد ما يجب على المستخدمين التعرض له وتزيل ما يعد مسيئا، وهذا كله بالاعتماد على البيانات.

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي ————— د/ عائشة كريكت

لذا وفي سبيل تأكيد ذلك تناولت الدراسة العناصر التالية: الإشكالية، الذكاء الاصطناعي نشأته وأهميته، خوارزميات التواصل الاجتماعي ووظائفها وأخيرا المجالات الأخلاقية التي تنظمها خوارزميات الذكاء الاصطناعي.

الإشكالية:

لقد طورت التكنولوجيا اليوم مختلف مجالات الحياة الإنسانية، وهذا من خلال توفير بدائل أساسها مساعدة البشر وتسريع وتيرة العمل والأداء، حيث لم يعد الانسان هو العامل المحوري في ذلك نظرا لتدخل الآلة في العديد من مهامه وكل هذا ربعا للوقت والجهد، حيث تشير الإحصائيات الصادرة عن مؤسسة Research and Market (2019) انه من المتوقع ان ينمو سوق الذكاء الاصطناعي العالمي من 28.42٪ مليار دولار في 2019 الى 40.74٪ دولار بمعدل نمو سنوي يبلغ 43.39٪ نتيجة جائحة كورونا، وتزايد الاعتماد على الآلات الذكية والروبوتات، ومن المتوقع ان يصل الى 400.2 مليار دولار بحلول عام 2027¹.

ويعد ذلك شكلا من أشكال استخدام ما يعرف بالذكاء الاصطناعي Artificial intelligence والذي يقوم على المعالجة الحاسوبية للمعلومات والبيانات، وإعطاء نتائج واتخاذ قرارات بطريقة مماثلة للبشر وهذا من خلال تطوير أنظمة قادرة على القيام بنفس العمليات التي يقوم بها الإنسان، ولعل خوارزميات مواقع التواصل الاجتماعي هي احد النماذج، حيث تقوم باختيار المحتوى عبر هذه المواقع بالاعتماد على البيانات التي تجمعها حول كل مستخدم لتضع له المحتوى الذي يتناسب مع اهتماماته، وهي بذلك تعمل عمل الصحفي او حارس البوابة في المؤسسة الإعلامية، فهي تتقصى دائما ما يهتم به المستخدم واي مضمون يتناسب مع اهتماماته.

وانطلاقا من ذلك يمكن القول انه قد يكون لها دور في أخلقة محتوى المنصات الافتراضية، وهذا من خلال حجب أو حذف المحتوى غير الأخلاقي عبرها، سواء كان خطاب كراهية وتمييز، أو محتوى عنف او إرهاب، الأخبار الكاذبة والمزيفة أو حتى محتوى ينتهك الخصوصية، خاصة في ظل عدم وجود قوانين واضحة تنظم نشاط المستخدم عبر هذه المنصات، وبالتالي يمكن أن تنظم هذه المواقع نفسها بنفسها من خلال الخوارزميات التي يمكنها ان تتحكم في المحتوى بالطرق المناسبة، وبالتالي يمكن ان توفر الحماية للمستخدمين من المحتويات غير الأخلاقية التي تغزو مواقع التواصل الاجتماعي خصوصا منها محتوى الكراهية والعنف.

وعليه فهذه الدراسة جاءت لتبحث في إمكانية التنظيم الآلي للمحتوى الرقمي بالاعتماد على خوارزميات الذكاء الاصطناعي، وهذا من خلال محاولة الإجابة عن الإشكالية التالية: هل يمكن ان يكون لخوارزميات الذكاء الاصطناعي دور في أخلقة محتوى مواقع التواصل الاجتماعي؟
بالإضافة إلى التساؤلات الفرعية التالية:

- ما هي وظائف الخوارزميات عبر مواقع التواصل الاجتماعي؟
- ما هي اهم المجالات الأخلاقية التي يمكن أن تنظمها الخوارزميات عبر مواقع التواصل الاجتماعي؟

أهمية الدراسة:

تتمثل أهمية هذه الدراسة في محاولة توضيح ما إذا كان هناك علاقة بين خوارزميات مواقع التواصل الاجتماعي وأخلاقيات النشر عبرها، وهذا من خلال إمكانية أن يكون لها دور إيجابي في الوقاية من المشكلات الأخلاقية التي ظهرت مع كثرة استخدام هذه المواقع، وتزايد نشاط الافراد بإنتاج المحتوى، حيث أصبح كل مستخدم بمثابة الصحفي او المؤسسة الإعلامية، يقدم الأخبار والمعلومات وي طرح الأفكار والآراء بكل حرية، وكل ذلك في ظل عدم معرفته بمبادئ وأخلاقيات النشر. الدراسة من بين الدراسات التي تبحث عن إجابات لمشكلة أخلاقيات مواقع التواصل الاجتماعي، والتي من خلالها يتم فتح آفاق جديدة للباحثين في التعمق أكثر في الموضوع، خاصة وان الإشكالات البحثية في دراسات مواقع التواصل الاجتماعي جديدة، وظاهرة الذكاء الاصطناعي من المواضيع التي لازالت في بداياتها خاصة في علاقتها بظواهر أخرى مثل الأخلاق.

أهداف الدراسة:

تتمحور أهداف الدراسة في:

- الكشف عما إذا كان هناك دور لخوارزميات مواقع التواصل الاجتماعي في تنظيم المحتويات المنشورة عبرها سواء كانت محتويات المستخدمين العاديين او محتوى المؤسسات الإعلامية، وهذا كوظيفة إضافية بالإضافة إلى وظائف أخرى أهمها: ترتيب المحتوى واقتراحه وفق اهتمامات المستخدم أي إدارة المحتوى
- البحث في علاقة الخوارزميات بمناهضة خطاب الكراهية والمضامين المتطرفة، التمييز العنصري، الأخبار الكاذبة وغيرها من المضامين للأخلاقية والضارة عبر هذه المواقع، وهذا كله في إطار التوجيه الأخلاقي لمحتوى المنصات الرقمية حتى لا يكون لها أضرار على المدى الطويل.

2. الذكاء الاصطناعي: مفهومه ونشأته

يمكن القول عن الذكاء الاصطناعي انه المعالجة الآلية للمعلومات مع الوصول الى نتائج مماثلة للعقل البشري، بحيث تتبنى الأنظمة الآلية طريقة التفكير البشري وتكون قادرة على معالجة المعلومات واتخاذ القرارات، وقد تعد بذلك عصرا بديلا للعديد من العمليات البشرية. يساعد الذكاء الاصطناعي وغيره من تكنولوجيات التحليلات المتقدمة في نقص تكاليف معالجة المعلومات، ويمكن للخوارزميات الحديثة القائمة على التعلم الآلي والبيانات الضخمة والحوسبة

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي _____ د/ عائشة كريكت

السحابية من جمع كميات هائلة من البيانات في ثوان عديدة وتحليلها لإنتاج أفكار دقيقة حول العمليات والسلوكيات بما يحفز الابتكار من اجل إحداث تغييرات أساسية في الإنتاجية والنمو.

وأطلق هذا الاسم لأول مرة في مؤتمر عقد في "دارتموث كوليدج هانوفر" في عام 1956، وقد نشأ المفهوم الأولي له بفضل عدة عوامل أهمها:

- 1- البيانات الضخمة: من خلال الكم الهائل للبيانات ومصادرها
- 2- وسائل التواصل الاجتماعي: ساهمت في سهولة تداول تطبيقات الذكاء الاصطناعي
- 3- البرامج والبيانات مفتوحة المصدر: تسمح بقضاء وقت اقل في البرمجة الروتينية²

3. الخوارزميات في عصر مواقع التواصل الاجتماعي

1.3. تعريف الخوارزمية:

الخوارزمية تأتي من اسم عالم الرياضيات في القرن التاسع "محمد بن موسى الخوارزمي"، وهي باللغة اللاتينية Algoritmi

وهي مجموعة من الإجراءات المرتبة ترتيبا منطقيًا والتي يتم تنفيذها للوصول الى هدف او ناتج مطلوب، وهي اختصار لكلمة: التعليمات البرمجية"، وهي التعليمات التي يكتبها مبرمج ويجمعها لاننتاج وحدة قابلة للتنفيذ، وتعرف أيضا باسم برنامج.

ولا يمكن لأنظمة الذكاء العمل بدون الخوارزميات³.

وعليه فان للخوارزميات قدرة على التحكم في بيانات ما، هذه البيانات قد تكون متعلقة بالمستخدم على مواقع التواصل الاجتماعي او بالمحتوى الذي ينتجه أو يتلقاه، وهذا ما يعرف بإدارة المحتوى وبالتالي قد تسمح هذه الوظيفة بإدارة المحتوى الضار على هذه الموقع من خلال انشاء خوارزميات خاصة بذلك.

2.3. تعريف مواقع التواصل الاجتماعي:

هي عبارة عن مواقع للاتصال بين الأفراد، بحيث تسمح بتبادل الرسائل (نصية، صوتية، او مصورة) بين مستخدميها، بالإضافة الى نشر محتوياتهم الشخصية او مشاركة محتويات الآخرين والتفاعل معها.

وقد تطورت الى منصات تستخدم في مختلف المجالات السياسية، الاقتصادية والأمنية وغيرها، وهذا بدوره أدى الى وجود مضامين متنوعة، منها ما هو مفيد وفي نفس الوقت ما هو غير ذلك.

وتعرف أيضا أنها مجموعة من تطبيقات الإنترنت التي تبنى على أسس إيدولوجية وتكنولوجية من الويب 2، والتي تسمح بإنشاء وتبادل المحتوى الذي يتم انشاءه بواسطة الإنترنت، ليتحول الاتصال الى حوار تفاعلي⁴.

3.3. وظائف الخوارزميات في مواقع التواصل الاجتماعي:

يمكن تصنيف الخوارزميات بحسب الوظائف الي تؤديها الي وظائف أساسية ووظائف مخصصة، وتتعلق الوظائف الأساسية بترتيب البيانات وفق نسق معين وتشمل: الترتيب بالإدراج، الترتيب النوعي، الترتيب بالفرز، الترتيب بالدمج، الترتيب السريع، الي جانب خوارزميات أخرى متعلقة بهيكله البيانات.

ويمكن تفصيل المعاني المرتبطة ببعض أهم وظائف الخوارزميات في الأدوار التالية:

1) وظيفة ترتيب نتائج البحث وفقا لكلمات مفتاحية:

تقوم هذه الخوارزمية بالتعامل مع كميات وفيرة من البيانات التي تنتجها محركات البحث وقواعد البيانات (Meta Data) بشكل انتقائي، بحيث تنتقي فقط النتائج المرتبطة مباشرة بالكلمة التي يبحث فيها المستخدم، وقد طور الباحثون عددا من الخوارزميات التي يمكنها القيام بهذه الوظيفة منها "مشروع كيفيتس" الذي يثبت قوة المنطق الذي تمثله أنظمة الخوارزميات، حيث يقومون بانتقاء مصادر البيانات ذات الثقة وفقا لتوصيات مختصين ثم ينظمونها في بناء ترابتي يمثل آلية لربط المحتوى الالكتروني بكلمات بحث محددة.

يمكن القول ان هذه الوظيفة هي أساس وجود الخوارزميات وأصل عملها، بحيث انها تختار للمستخدم نوعه المفضل من المضامين بناء على اهتماماته الي يبحث عنها عبر مواقع التواصل الاجتماعي (يوتيوب مثلا) وهذا في الوقت الذي يظن فيه المستخدم انه من يختار ذلك، وبالتالي قد تكون خيارات المستخدم عبر المنصات الرقمية هي خيارات موجهة وليست خيارات ذاتية، وبما أن هذه الخوارزميات قادرة على اختيار المحتوى المناسب للمستخدم بالاعتماد على بياناته فإنه وفي نفس الوقت يمكن لها ان تزيل ما تعتبره غير مناسب او ضار.

2) خوارزميات الإرهاب: ان الهدف الأساسي لهذه الخوارزميات هو تتبع أنشطة الشخصيات او المجموعات ذات الميول الإرهابية عبر شبكات التواصل الاجتماعي من خلال الصفحات المشبوهة والمنشورات التي تستهدف حشد الأنصار او تجنيد المقاتلين للتنبؤ بأي عمليات إرهابية محتملة.

ومثال ذلك ما قامت به روسيا من توظيف لهذه الخوارزمية عام 2015 على شبكتها الخاصة بالتواصل الاجتماعي "Vkontakte" والتي تضم أكثر من نصف مليار مستخدم، وضبطت من خلالها ما يقارب 196 مجموعة داعمة لتنظيم الدولة والقاعدة، وما يفوق 180 ألف حساب لشخص مشارك بتنظيمات يمكن تقييمها بكونها ذات نشاط إرهابي⁵ (السيد، 2020)

إن هذا النوع من الخوارزميات يساعد في تصفية مواقع التواصل الاجتماعي من المحتوى الإرهابي، ويجنب المستخدمين التعرض للأفكار والمعتقدات المتطرفة خاصة وانه تقل الرقابة عبر هذه المواقع،

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي ————— د/ عائشة كريكت
ما يجعل المجال مفتوحاً أمام الإرهابيين، وعليه فالذكاء الاصطناعي من خلال الخوارزميات كفيل
بمواجهة ذلك.

4.3. الخوارزميات وإدارة المحتوى عبر مواقع التواصل الاجتماعي:

يوضح استشاري الاعلام الاجتماعي " خالد صافي" ان موقع فيسبوك يعتمد خوارزمية معقدة
لعرض المنشورات للأصدقاء او المتابعين للصفحات الرئيسية، فالصداقة بين اثنين ليست شرطا
لظهور المنشورات لكليهما فهناك مجموعة من المعايير كالاهتمامات المشتركة والفوارق الجغرافية
والديمغرافية، والخوارزميات تتقصى دائما ما يهتم به المستخدم عبر هذه المنصات وتعرض المناسب
لاهتماماته⁶.

فأغلب المحتوى المنشور عبر هذه المواقع هو من اختيار برامج حاسوبية تابعة للشركات المالكة
ويديرها مختصون في ذلك، بالاعتماد على بيانات المستخدمين وهذا ما كشفه تعليق حسابات
الرئيس الأمريكي السابق "دونالد ترامب"، ما يثير مخاوف حول الترويج للمحتوى الضار.
وفي سياق التعرف على مدى تأثير خوارزميات فيسبوك على بيانات المستخدمين الشخصية، أجرى
مركز Bio للأبحاث مسحا نشر في يناير 2019، توصل الى ان فيسبوك يستخدم خوارزميات تقوم
بتصنيف المستخدمين طبقا لما يسمى " الانتماء متعدد الثقافات " Multicultural affinity، وتتم من
خلاله وضع المستخدمين في قوائم تبعا لميولهم المعرفية والقومية بدلا من انتماءاتهم الفعلية.
وبالنظر الى خوارزميات يوتيوب، تهتم أيضا منصة فيسبوك بتقديم محتوى مشخص، بحيث تختار
المقاطع المفضلة للمستخدم وتضعها مباشرة على شاشته عند الاستخدام التالي.

اما خوارزميات تويتر فمنذ 2014 بدأت المنصة بتقديم اقتراحات متابعة معينة او الاطلاع على
تغريدات متداولة على نطاق واسع، وأضافت خلال السنوات التالية خوارزميات جديدة مثل: في حال لم
تشاهدها في حينها (In case you misse dit)، والخوارزميات التي يستخدمها تويتر الآن تشمل: Top
7Happening Now ، trends for you

ويرى "مارك زوكربيرغ" ان الذكاء الاصطناعي هو الحل لإدارة المحتوى، بحيث يسمح بإزالة
الحسابات الوهمية ومعالجة محتوى مليون مستخدم بعدة لغات⁸.

مثال: خوارزميات محركات البحث

تعد خوارزميات محركات البحث هي العمود الفقري للنجاح، لأنها تسمح بالبحث في ملايين
مواقع الإنترنت والعثور على المحتوى المناسب، والتعرف على الأخطاء الإملائية مع تقديم اقتراحات
للبحث في كلمات أخرى، لتحديد في النهاية المعلومات التي يجب تقديمها.

وعلى سبيل المثال تركز شركة Netflix على عرض محتوى مناسب لمشاهديها وبناء توصيات
بناء على السلوك البحثي على منصتها الرقمية على منصتها الرقمية، وتشير الشركة انها توفر ما

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي ————— د/ عائشة كريكت
يقارب من مليار دولار امريكي سنويا بفضل تقنية الذكاء الاصطناعي على آلية تدفقات المحتوى
والتفاعل مع العملاء.⁹

وما يؤكد قدرة الخوارزميات على إدارة أي محتوى، أيضا ما قامت به مواقع التواصل الاجتماعي بحق
القضية الفلسطينية ضمن فضاء الاعلام الرقمي، نتيجة تحكمها بالسياسة الإعلامية والخطاب السائد،
فخوارزميات تطبيقات فيسبوك وانستغرام مبرمجة لتزيل أوتوماتيكيا المحتوى المناصر لفلسطين،
في المقابل الناشطين المناصرين لها لجأوا الى آليات مؤثرة مثل تخفيض تقييم فيسبوك على
Appstore و Google play store¹⁰.

وعليه فبامتلاك الخوارزميات القدرة على إدارة أي محتوى عبر مواقع التواصل الاجتماعي فبالتالي
يمكن لها ان تدير المحتوى غير الأخلاقي سواء المتضمن في المنشورات النصية، او المنشورات
الصوتية والمصورة من خلال تشخيص المحتوى بناء على مجموعة من المعايير التي نجدها أحيانا
ضمن سياسة الاستخدام عبر هذه المواقع، ويكون التحكم به عن طريق الحذف التلقائي للمحتوى.

5.3. خوارزميات الذكاء الاصطناعي و تحديد المحتويات السامة

شهدت السنوات الأخيرة انتشار متسارع على شبكة الإنترنت للغة المسيئة، ففي بعض الأحيان
أدت التعليقات السامة على شبكة الإنترنت إلى إثارة العنف على أرض الواقع، لذا فقد استعان العديد
من شركات التكنولوجيا العملاقة بالخوارزميات في عملية المراجعة للمحتوى لتحديد السام منه، وهذا
ما قامت به شركة "جيجسو" التابعة لـ Google حيث اقترحت نموذج لتصنيف التعليقات السامة تحت
وسوم: "سام"، "تهديدي"، "مهين"، إضافة الى وجود مكتبة "Detoxify" لتحديد المضامين المؤذية حتى
يتم معرفتها من طرف المستخدمين.

فمثلا جملة " سمتم من كتابة هذا المقال الغبي " تسجل 99.7% على مقياس السمية، في حين يؤدي
حذف كلمة الغبي " منها الى خفض هذه النسبة الى 0.05%¹¹.

لذا فإن وجود خوارزميات الذكاء الاصطناعي التي من شأنها ان تتعرف على المحتوى الضار عبر
مواقع التواصل الاجتماعي هو امر في غاية الأهمية، ذلك لان الخطابات المسيئة سواء من خلال
المنشورات او التعليقات قد انتشرت بشكل رهيب لدرجة انه قد أصبحت هذه المنصات فضاءات لتبادل
التعليقات المسيئة والمضايقات، فالكشف الآلي قد يحل محل المراقبين البشر رغم انه لا يزال أمامه
طريق طويل، لذا وفي سبيل تنظيم المحتوى الرقمي فانه لازال هناك الكثير امام الذكاء الاصطناعي
لكشف عن المحتوى السام او الضار.

4. المجالات الأخلاقية التي تنظمها الخوارزميات عبر مواقع التواصل الاجتماعي

1.4. مفهوم الأخلاقيات:

الأخلاقيات (Ethique) تعود الى الكلمة اليونانية Ethos التي تعني العادات الأخلاقية¹² وحسب "سمير الباجوري" هي نظام من المبادئ المبنية على القيم الأخلاقية التي توجه عمل ما¹³ ولما كانت الأخلاقيات المهنية تمثل مجموعة من القيم المتعلقة بالممارسة اليومية للصحفيين وفي نفس الوقت تعمل على حماية الجمهور من أي استخدام غير مسؤول¹⁴، فإن أخلقة المحتوى عبر مواقع التواصل الاجتماعي تتمثل أساسا في نشر مضامين تلتزم بمختلف المبادئ والقواعد التي تنص عليها المواثيق الأخلاقية في مجال الإعلام او الخاصة بمواقع التواصل الاجتماعي عموما، وهذا لحماية حقوق الأفراد والجماعات عبر هذه المنصات.

ومن اهم القيم الأخلاقية المتفق عليها للنشر عبر شبكات التواصل الاجتماعي حسب دراسة بعنوان: "أخلاقيات استخدام شبكات التواصل الاجتماعي من وجهة نظر مدرسي الإعلام والقانون في الجامعات الأردنية": الأمانة، المصداقية، المحافظة على الآداب العامة، الدقة، تجنب السب أو الذم، التبادل الثقافي¹⁵.

2.4. مفهوم أخلاقيات مواقع التواصل الاجتماعي:

عند تناول أخلاقيات الإعلام والمسؤولية الاجتماعية فقد كان التركيز على الإعلامي لأنه مستخدم وسائل الإعلام، لكن مع ظهور الإعلام الجديد خاصة مواقع التواصل الاجتماعي فقد أصبح هناك حاجة للحديث عن أخلاقيات ومسؤولية اجتماعية للمواطن العادي، حيث يرى بعض الباحثين انه من الضروري وضع أخلاقيات جديدة تناسب هذه المستجدات، في حين يرى البعض انه يمكن تنقيح وتعديل الأخلاقيات الخاصة بالإعلام التقليدي فتشمل الإعلام الجديد ومواقع التواصل الاجتماعي.

وتوجد تحديات كثيرة لوضع اخلاقيات وضوابط للإعلام الجديد والذي تتضمنه مواقع التواصل الاجتماعي مثل الاستخدام واسع النطاق في كل أنحاء العالم مما يجعل من الصعب وضع أخلاقيات وضوابط تناسب كل المستخدمين باختلاف الثقافات والحضارات، وتوجد إشكالية أخرى في إمكانية احترام الخصوصية وفي نفس الوقت تجنب سلبيات إخفاء الهوية الشخصية.

وتتمثل أكبر المشاكل الأخلاقية لمواقع التواصل الاجتماعي خاصة موقع الفيسبوك في التحرش والتهديدات وانتهاك الخصوصية وتعريض الحياة الخاصة للعامة وغيرها

وقد أوضحت دراسة تابعة للأمم المتحدة عام 2007 بضرورة وضع ميثاق أخلاقي ليكون مرجعية عالمية ودولية للأخلاقيات في مجتمع تكنولوجيا المعلومات، بالإضافة الى أهمية التواصل مع الخبراء لتطوير وابتكار الطرق التكنولوجية لدعم الإيجابيات وتجنب السلبيات خاصة مع ما يرتبط بالحفاظ على الخصوصية وحماية المعلومات، وتعاني معظم الدول من التجاوزات الأخلاقية التي يقع فيها مستخدمو الاعلام الجديد خاصة مواقع التواصل الاجتماعي في ظل غياب الوعي وعدم تطبيق الاخلاقيات والمواثيق الا في حالات قليلة جدا¹⁶.

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي _____ د/ عائشة كريكت

ومن اهم المجالات الأخلاقية التي تنظم عبر مواقع التواصل الاجتماعي مايلي:

3.4. الخوارزميات وإدارة المحتوى الكاذب والمضلل:

الأخبار الكاذبة او الزائفة تطرح محتوى مختلف او معلومات خاطئة او نظريات مؤامرة بشكل مباشر، وهي تمثل ضررا بالتأكيد، وهناك ما يعرف بالتزييف العميق الذي يعني " التقول على شخص بشيء لم يقله" من خلال فيديوهات مزيفة لشخصيات تحمل نفس الصوت وتعابير الوجه. يستخدم موقع فيسبوك الذكاء الاصطناعي في التصدي للأخبار المزيفة Fake news عبر استخدام التدقيق البصري والبحث العكسي عن الصور، وتحليل بياناتها الوصفية مثل اين تم التقاط الصورة او الفيديو لكن تواجه صعوبات في تحليل الصور الساخرة او مقاطع الفيديو بسبب التلاعب بالألفاظ والاختلافات اللغوية والثقافية.

ويستخدم متقصد الحقائق تطبيقات الذكاء الاصطناعي في اتخاذ القرار في ما يجب التحقق منه يوميا، حيث يتم اجراء فحص تلقائي RobocheKing بالاعتماد على الربط بقواعد بيانات المعلومات ، فعلى سبيل المثال إذا قال احد الأشخاص " انخفضت العمالة بنسبة 10٪ منذ عام 2016 يقوم نظام الذكاء الاصطناعي بالعثور على الأرقام الصحيحة، وانشاء رسم بياني والاجابة عنها في وقت وجيز، فبينما تستغل المعلومات المضللة تقنيات جديدة تستمر في العمل على نطاق واسع، تكون تقنيات الذكاء الاصطناعي أدوات رئيسة لعمل غرف الأخبار مثل كشف " التزييفات العميقة" والتحقق من صحة الصور، الفيديو.

وأیضا طور باحثون من جامعة "واترلو" الكندية نظام قائم على الذكاء الاصطناعي للتحقق من صحة المعلومات، بعد الاطلاع عليها يقدم مؤشر بصحتها ام لا، وقد نجح 9 مرات من أصل 10. واستخدم تويتر وسوم للإبلاغ عن المحتوى المتلاعب به وكان ذلك على محتوى انشأه مدير وسائل التواصل الاجتماعي بالبيت الأبيض، كم ان "يوتيوب" قلل الترويج لمقاطع فيديو المؤامرة، بعد ان غيرت الشركة خوارزميات توصياتها، وهذا حسب دراسة في جامعة كاليفورنيا سنة 2020. ان قيام مواقع التواصل الاجتماعي بالكشف عن المحتوى الكاذب والأخبار الخاطئة من خلال الخوارزميات من شأنه ان يحد من كمية الأخبار الكاذبة عبرها، خاصة وان هذه المنصات ينشر فيها الجميع دون مراعاة لمصادر موثوقة خاصة من قبل الأشخاص، حيث ان الخبر الكاذب قد يؤدي الى تغليب الرأي العام وهذا في الغالب ما يحدث خاصة في بعض القضايا المصيرية التي يتابعها الرأي العام باهتمام¹⁷.

وبالتالي فهذه الخوارزميات تؤدي دور المصفاة لكل ما هو مزيف او مضلل، رغم انه لازالت هناك حاجة لمعرفة المزيد عن كيفية عمل هذه الخوارزميات في الكشف عن الاخبار الكاذبة، وحاجة أيضا لتطوير خوارزميات جديدة تعمل خصيصا في هذا المجال.

4.4. الخوارزميات وإدارة المحتوى المتطرف والإرهاب:

ان منصات التواصل الاجتماعي يمكن ان تعزز المحتويات المتطرفة من خلال نظام التوصية، حيث تساعد أنظمة التوصية في العثور على المضمون التالي بناء على العناصر التي سبق استخدامها، فمثلا في اليوتيوب يتم التوصية بمضامين مماثلة، فمثلا اذا تم مشاهدة فيديو لمحتوى متطرف يقوم باقتراح فيديو مماثل، ومن جهة أخرى فان فيسبوك بدوره تعرض لانتقادات نتيجة عدم إيقافه البث الفوري لفيديو الهجوم الإرهابي على مسجد " كرايستشيرش" في نيوزيلندا، حيث كان بثا مباشرا نشره صاحب الهجوم، وفي تويتر ويوتيوب أيضا لم يتم إزالته الا بعد مدة، وهي بذلك قد فتحت المجال أمام المتطرفين لاستغلال منصاتهما في أعمال غير إنسانية.

أطلق فيسبوك خدمة "ريدايركت إينشياتيف" لمكافحة التطرف من خلال توجيه الأشخاص المستخدمين للكلمات العنيفة والمتطرفة الى مجموعات توعوية، وعلى سبيل المثال مجموعة " لايف أفتر هايت" التي جاءت كرد على حركات اليمين المتطرف المناهية بتفوق البيض في الولايات المتحدة الأمريكية.

ويعمل فيسبوك حاليا على رقائ مضممة لتصفية محتوى الفيديو المباشر، حيث انه بمجرد تحديد ما إذا كان محتوى متطرف او منتهك لسياسة الموقع يتم حذفه تلقائيا، وخلال فترة زمنية قصيرة، ويعمل في فيسبوك حاليا 15000 موظف يتصفحون المشاركات لإزالة أي محتوى مسيء¹⁸

واكد "روزون" ان خوارزميات الذكاء الاصطناعي تراقب جميع المشاركات بلغاتها المختلفة سواء الاسبانية او العربية بجانب الإنجليزية تحقيقا لأقصى درجات الحماية للمستخدمين (ممدوح، 2020). كما ان تويتر قد وضعت رابط " طلب المساعدة" للأشخاص الذين تعرضوا شخصا لمحتويات متطرفة¹⁹.

وبعد حادثة الهجوم الإرهابي في نيوزيلندا أقامت حكومة نيوزيلندا اتفاقا مع شركات وسائل التواصل الاجتماعي للقضاء على المحتوى الإرهابي، وهذا بدعوتها الى تعزيز المساءلة والشفافية من جهة، ومن خلال مراجعة المحتوى وتقييمه من جهة أخرى، بحيث تولى أهمية لعمليات إدارة المحتوى، وهذا كله يأتي في إطار غياب هيئات مسؤولة عن مكافحة التطرف عبر الإنترنت²⁰.

5.4. الخوارزميات وإدارة محتوى الكراهية:

ان خطاب الكراهية هو الخطاب المتعمد لغاية نشر او ترويج او حتى تبرير الكراهية تجاه جماعة عرقية او قومية، او أي شكل من أشكال الكراهية المبنية على التعصب، وكره الأجانب او معاداة السامية وغير ذلك من اشكال الكراهية المبنية على التعصب والتمييز والعداء للأقليات والمهاجرين. وفي هذا الإطار أكدت دراسة وجود علاقة بين التغريدات والمنشورات المتعلقة باللاجئين في تويتر وفيسبوك وبين الاعتداءات عليهم في نفس الفترة، وبحسب الدراسة فإنه كلما زاد عدد التغريدات

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي ————— د/ عائشة كريكت والمنشورات التي وردت فيها كلمة لاجئين flykting ارتفع عدد الحرائق التي يفتعلها متطرفون في مساكن يقطنها لاجئون.

وتوضح الدراسة ان خوارزميات مواقع التواصل الاجتماعي تساهم في خلق ما يعرف بـ "غرف الصدى" (Echochambers)، وهو مصطلح يشير الى تعرض الأفراد للمحتوى الذي يتوافق مع تفضيلاتهم الشخصية أكثر من أي محتوى آخر، وبعبارة أخرى فإن هذه الخوارزميات قد تزيد من فرص تعرض الأشخاص ذوي النزعة العنصرية إلى محتوى إعلامي يصور المهاجرين واللاجئين كخطر على المجتمع²¹.

يحمل القانون شركة الوسائط الاجتماعية المسؤولية إذا تم استخدام خوارزميتها للتضخيم او التوصية بمحتوى ذي صلة مباشرة بقضية تنطوي على تدخل في الحقوق المدنية، وفي القضايا المتعلقة بأعمال الإرهاب الدولي، بينما ينشر فيسبوك بانتظام جهوده المستمرة لإزالة منشورات مجموعات الكراهية وتضييق الخناق على تنسيقها باستخدام منصته، حتى التقارير الداخلية للشركة تجادل بأنها لم تفعل ما يكفي تقريبا لوقف موجة التطرف على الموقع²².

وتعتمد فعالية الخوارزميات في كشف المواد الدعائية للكراهية على جودة البيانات ومدى توفرها. ومما يثبت قدرتها على إدارة محتوى الكراهية لكن في الاتجاه السلبي لا الإيجابي هو الحذف التلقائي لكلمات متعلقة بمواقع فلسطينية مثل: الأقصى وكل ما يدور في فلكه مثل: هاشتاغ الأقصى وكان التبرير خطأ تقني، حيث مارست نوعا من التمييز ضد المحتوى الفلسطيني مع انه ليس محتوى كراهية²³.

وقد اختبر الباحثون خدمتين تجاريتين شهيرتين: الأولى Perspective API التابعة لجوجل جيجسو والثانية Sift Ninja، تسمح كليهما للعملاء بالإبلاغ عن المحتوى المخالف في المنشورات او التعليقات من خلال جمل نموذجية مثل: "انا أكره (هوية ما)" او "انت مجرد (إهانة) بالنسبة لي"، وتقوم هذه الخدمة بالإبلاغ عن المشاركات والتعليقات وتحديد مدى أولويتها للمراجعة بناء على مقياس السمية الخاص بها.

وتعمل " جيجسو" الآن على تطوير ميزة تقوم على ترتيب المشاركات والتعليقات بناء على درجة الشك Perspective في المحتوى الذي يحرض على الكراهية وتحويله الى مراقبين²⁴.

6.4. الخوارزميات والمحتوى الإباحي والعنيف:

في بداية استخدام فيسبوك كان حظر الإباحية يعتمد على بلاغات المستخدمين، حيث يتم حظر المنشورات التي تتضمن إباحية او عنف بعد مراجعتها من طرف مراقبين، لكن مع تجاوز عدد المستخدمين للمليارين، تم الاستعانة بالذكاء الاصطناعي في حجب المحتوى الإباحي والعنيف من خلال تحليل الصور ومقاطع الفيديو التي من المحتمل ان تتضمن ذلك من طرف الخوارزميات²⁵.

خوارزميات الذكاء الاصطناعي وأخلاقيات محتوى مواقع التواصل الاجتماعي ————— د/ عائشة كريكت
ويوتيوب نموذج لذلك حيث تستخدم تقنيات التعلم الآلي لتحليل محتوى ملايين المقاطع على الموقع
وتشمل المحتوى العنيف، ومن جهتها شركة فيسبوك تستخدم الخوارزميات في إزالة المنشورات
المسيئة، وأحياناً بسرعة كبيرة لدرجة ان المستخدم قد لا يراه²⁶.

خاتمة

نستنتج مما تقدم ان خوارزميات الذكاء الاصطناعي لها دور فعال في أخلقة محتوى مواقع التواصل
الاجتماعي وهذا إذا تم استخدامها في الاتجاه الصحيح والإيجابي بحيث ان العلاقة بينهما هي علاقة
وظيفية، فبالإضافة إلى دورها الأساسي والمرتبط بتحديد المضامين حسب اهتمامات المستخدم
وتفضيلاته، فإنه يمكن ان يكون لها وظائف أخرى تتعلق بإدارة المحتوى عبر هذه المواقع خاصة منه
غير الأخلاقي والضار.

وهذا الدور يتمحور حول المجالات التالية:

- 1- إدارة المحتوى الكاذب من خلال: تعليق الحسابات، التدقيق البصري، البحث العكسي عن الصور
وتحليل بياناتها الوصفية، الربط بقواعد البيانات وغيرها
- 2- إدارة المحتوى الإرهابي والمتطرف من خلال: حجب المنشورات، خدمات مثل " ريدايركت
اينيشياتيف"، الحذف التلقائي للمحتوى بالاعتماد على رقائق للتحليل والتصفية
- 3- إدارة محتوى الكراهية من خلال: تدريب الخوارزمية بالاستعانة بالبيانات، الحذف التلقائي
لل كلمات أو العبارات التي تحض على الكراهية، خدمة الإبلاغ عن محتوى الكراهية في
المنشورات أو التعليقات بالاعتماد على الجمل النموذجية، مقياس السمية الذي يحدد النسبة
المئوية للإساءة باستخدام الكلمات او العبارات
- 4- إدارة محتوى الإباحية والعنف من خلال: تحليل الصور ومقاطع الفيديو، الإزالة الفورية
للمحتوى الإباحي او العنيف فور مراجعته من طرف مشرفي المحتوى.

الهوامش

¹ عبد الحميد، عمرو محمود، محمد (2020). توظيف تطبيقات الذكاء الاصطناعي في إنتاج المحتوى الإعلامي وعلاقتها
بمصادقته لدى الجمهور المصري. مجلة البحوث الإعلامية. العدد (55)، ص 2817.

² مؤسسة آل مكتوم، محمد بن راشد للمعرفة. (2019)، استشراف مستقبل المعرفة. دبي-الامارات العربية المتحدة.
الغدير للطباعة والنشر، ص 7-9.

³ موسى، عبد الله و بلال، احمد حبيب، (2019)، الذكاء الاصطناعي، الطبعة الأولى، القاهرة، المجموعة العربية
للتدريب والنشر، ص 98-99.

⁴ عبد الفتاح، علي، (2020)، الاعلام الاجتماعي، عمان، دار اليازوري للنشر والتوزيع، ص 4.

- ²⁰ عبد الحميد، عمرو محمود، محمد، مرجع سبق ذكره، ص33-34.
- ²¹ مرعي، زهرة، مرجع سابق
- ²² جمال، منة الله. (2021)، كيف تساعد خوارزميات توصية وسائل التواصل الاجتماعي في نشر الكراهية، الرابط <https://alwafd.news/%D8%AA%D9%83%D9%86%D9%88%D9%84%D9%88%D8%AC%D9> تم التصفح يوم 5 سبتمبر 2021، الساعة 14:28.
- ²³ مرعي، زهرة، مرجع سابق
- ²⁴ هاو، كارين. (2021)، هل ينجح الذكاء الاصطناعي حقا في مراقبة خطاب الكراهية. الرابط <https://technologyreview.ae>، تم التصفح يوم 4 سبتمبر 2021، الساعة 17:14.
- ²⁵ الراوي، طه. (2019). دور الذكاء الاصطناعي في تربع فيسبوك على عرش شبكات التواصل. الرابط <https://www.noonpost.com/content/34941>، تم التصفح يوم 3 سبتمبر 2021، الساعة 17:18.
- ²⁶ منظمة هيومن رايتس ووتش. (2021)، منصات التواصل الاجتماعي تزيل أدلة جرائم الحرب، متوفر على الرابط <https://www.hrw.org/ar>، تم التصفح يوم 4 سبتمبر 2021، الساعة 14:15.