تحديات الذكاء الاصطناعي في تحليل المحتوى الرقمي العربي في البحوث الاجتماعية

The challenges of artificial intelligence in analyzing digital Arabic content in social research

معمري عبدالوهاب★ [1]، مشيرة حسن العشيري[2]، سليماني حليمة[3]

[1] جامعة طاهري محمد بشار، (الجزائر)، mammeri.abdelouahab@univ-bechar.dz

[2]جامعة دمياط، (مصر)، moshera_elashry@yahoo.com

[3] جامعة أبي بكر بلقايد – تلمسان (الجزائر)، hmammeri08@gmail.com

تاريخ الإستلام: .../.../..... تاريخ قبول النشر: .../.../..... تاريخ النشر: ../.../.....

ملخص:

تهدف الدراسة فهم التحديات، التي تواجه الذكاء الاصطناعي، في تحليل المحتوى العربي الرقمي، في البحوث الاجتماعية، بالمنهج الوصفي والمنهجين الاستنباطي والتحليلي، ومراجعة العديد من الدراسات التي اهتمت بالمحتوى الرقمي العربي، وكذا الذكاء الاصطناعي، تجلى لنا، بأنه مع ثورة التكنولوجيا، وتطور تقنيات الذكاء الاصطناعي، تستخدم تلكم التطبيقات، في تحليل المحتوى عموما، وبشكل ملفت المحتوى الرقمي العربي، يحظى هو الآخر باهتمام واضح، ولكن المهمة معقدة، لصعوبة اللغة العربية، من حيث تواجد المحتوى العربي، رقميًا وحضوره في الأنترنت، ومن حيث التصنيف والتحليل، لوجود تحديات كثيرة، منها ما هو منهجي أو اجتماعي، أو تقني.

الكلمات مفتاحية: المحتوى الرقمي العربي؛ الذكاء الاصطناعي؛ التعلم الآلي؛ تحليل المحتوى؛ تصنيف النص.

Abstract:

The study aims to understand the challenges facing artificial intelligence in analyzing digital Arabic content, in social research, with the descriptive approach and the deductive and analytical approaches, and to review many studies that focused on Arab digital content, as well as artificial intelligence, it became clear to us, that with the revolution of technology, and the development of technologies Artificial intelligence, using these applications, in analyzing content in general, and in a striking way, Arab digital content, also receives clear attention, but the task is complex, due to the difficulty of the Arabic language, in terms of the presence of Arabic content, digitally and its presence on the Internet, and in terms of classification and analysis, due to the presence of challenges Many, including what is systematic, social, or technical.

★معمري عبدالوهاب، جامعة طاهري محمد بشار، mammeri.abdelouahab@univ-bechar.dz .

Keywords: Arabic digital content; Artificial intelligence; machine learning; content analysis; text classification.

1 مقدمة:

اليوم يلعب الذكاء الاصطناعي دورا غير مرئي في كثير من الأحيان في الحياة اليومية، حيث تعمل على تشغيل محركات البحث وتوصيات المنتجات وأنظمة التعرف على الكلام، هناك الكثير من الضجيج حول تطوير الذكاء الاصطناعي، وهو أمر متوقع من أي تقنية ناشئة تظهر، فهو يستفيد من أجهزة الكمبيوتر والآلات لمحاكاة قدرات العقل البشري على حل المشكلات واتخاذ القرارات واستعماله في مختلف وتفاصيل تحركاتنا في هذا العصر الرقمي.

تعاظمت الكميات الهائلة من المحتوى الذي ينشئه المستخدمون والمتاح في مختلف المنتديات والمدونات ومواقع التواصل الاجتماعي والبحوث عبر الإنترنت ولدى المكتبات الرقمية. يحتوي هذا المحتوى على معلومات حول مواقف الأفراد وتصوراتهم وآرائهم تجاه مختلف المنتجات والخدمات والأحداث الاجتماعية والقضايا السياسية. يتيح المحتوى الغني بالرأي اكتشاف معرفة جديدة في مجالات مختلفة مثل الرعاية الصحية والتعليم والسياسة والأمن. فهذا التزايد الكبير جعل البحث الفعال والتحليل واتخاذ القرارات أكثر صعوبة، للأشخاص المهتمين بهذا النوع من مصادر البيانات الغنية مثل الباحثين في مجال الكمبيوتر والمعلومات والباحثين في المعلوماتية الأمنية وعلماء الاجتماع والوكالات الحكومية وعامة الناس. ولمواجهة هذه التحديات، من شأن الأدوات المتقدمة والآلية أن توفر استجابة فعالة. كما أشار كيم وآخرون، فإن البحث عن محتوى وسائل التواصل الاجتماعي وتحليله أمر شاق ويستغرق وقتا طويلا. ومع ذلك، فإن البحث غير الكافي عن محتوى وسائل التواصل الاجتماعي ينتج معلومات غير كاملة يمكن أن تضلل بدورها عملية صنع القرار من قبل الأفراد أو المنظمات على حد سواء.

بل ظهرت مجموعة من التحديات في تطوير أدوات دعم البحث والتحليل الآلي. على سبيل المثال، يعد جمع ودمج المحتوى الذي ينشئه المستخدمون، عادة ما يكون شبه منظم أو غير منظم، عبر مواقع التواصل الاجتماعي المختلفة تحديا، أن الافتقار إلى التنسيقات المشتركة في المحتوى الذي ينشئه المستخدمون يمثل مشكلة أيضا، بالمقارنة مع الموارد الأخرى عبر الإنترنت، مثل مستودعات الأخبار أو المقالات العلمية، غالبا ما يفتقر المحتوى الذي ينشئه المستخدمون إلى تنظيم هيكلي متسق، وبالتالي عند البحث عن محتوى ذي صلة يتعلق بموضوع أو حدث معين، مثل التعليقات المنشورة التي تمثل الآراء السائدة، غالبا ما يحتاج الأشخاص إلى زيارة مواقع وسائط اجتماعية متعددة لتصفح المحتوى ذي الصلة ودمجه مع الدعم البدائي إن وجد، عادة ما تختلف مواقع التواصل الاجتماعي الشائعة في تنظيمها أو هيكلها، مما يزيد من تقييد استخدام إجراء وصول واحد لجمع المحتوى عبر المواقع الإلكترونية.

تقدم اللغة بعدا آخر من أبعاد التعقيد. عادة ما يتم إنشاء المحتوى المتاح في العديد من المواقع بلغات مختلفة. وفقا لإحصاءات الإنترنت العالمية، فإن أكثر من 70٪ من مستخدمي الإنترنت غير ناطقين باللغة الإنجليزية (Internet World Stats, 2020). ونتيجة لذلك، يجب أن يعالج الدعم الفعال للبحث والتحليل بشكل صحيح المشكلات المحيطة بالمحتوى متعدد اللغات الذي ينشئه المستخدمون. تقدم معظم تقنيات البحث والأدوات التحليلية الحالية أدوات مساعدة محدودة لجمع ودمج المحتوى الذي تم إنشاؤه بلغات مختلفة، ويرجع ذلك جزئيا إلى أنها مصممة لهياكل نحوية محددة أو تنسيقات تنظيم المستندات . لدعم اتخاذ القرارات المحسنة، يجب أن تسمح الأدوات الآلية للأشخاص بجمع وتخزين وتحليل المحتوى الضخم الذي ينشئه المستخدمون بلغات مختلفة، وتوفير المرونة المرغوبة والوصول في الوقت الفعلي تقريبا، وتوفير تصميمات عرض النتائج المناسبة.

أما عن لغتنا العربية فتتسم بقدرتها على استيعاب العلوم الحديثة، وتلبية أغراض الاتصال في الحياة، دون جعلها حكرا على أغراض محددة، وهو ما يحتم على الدارسين ضرورة مناقشة هذا الموضوع، ودعم ما يبذل من جهود لاستخدام التقنيات الحديثة بالأبحاث، والدراسات المتتالية، والعميقة، بدأت الدول العربية تمضي قدما في توسيع نطاق الاتصالات بالشبكة العنكبوتية إذ نبه بعض الخبراء إلى أن سنة:2018، عرفت في العالم العربي نحو:226مليون مستخدم للانترنت، بما يعادل أكثر من:55بالمائة من السكان، وذلك بمعدل يزيد على 7 بالمائة من المتوسط العلمي.

فجاءت هذه الدراسة لمعالجة الاشكالية التي فحواها: ما أهم التحديات التي تواجه الذكاء الإصطناعي في تحليل المحتوى الرقمي العربي في البحوث الإجتماعية؟

المنهج المتبع: من خلال المنهج الوصفي الذي اعتمدناه نحاول سرد مختلف الجوانب النظرية لمتغيرات الدراسة، كما نعتمد المنهج الاستنباطي والتحليلي لغرض تحليل مختلف التحديات التي توصل لها مطوري التطبيقات الرقمية في تعاملهم مع المحتوى الرقمي.

2 الأدبيات النظرية للذكاء الإصطناعي وتحليل المحتوى الرقمي العربي

1.2 الذكاء الإصطناعي:

1.1.2 ماهية الذكاء الإصطناعي:

ربما يكون هناك العديد من التعاريف الذكاء الاصطناعي كما أن هناك تفرعات كثيرة ضمن الذكاء الاصطناعي وتستمر الأرقام في الازدياد، ولكن نحاول سرد البعض منها.

يشير الذكاء الاصطناعي إلى "الروبوتات وأجهزة الكمبيوتر والآلات الأخرى ذات القدرة البشرية على التفكير وحل المشكلات" (McPherson, 2018, p. 4).

الذكاء الاصطناعي هو نظرية وتطوير أنظمة الكمبيوتر القادرة على أداء المهام التي تتطلب عادة الذكاء البشري، مثل الإدراك البصري، والتعرف على الكلام، واتخاذ القرارات، والترجمة بين اللغات.

يتيح الذكاء الاصطناعي للآلات التعلم من التجربة، والتكيف مع المدخلات الجديدة، وأداء مهام تشبه الإنسان. تعتمد معظم الأمثلة الذكاء الاصطناعي التي تسمع عنها اليوم – من أجهزة الكمبيوتر التي تلعب الشطرنج إلى السيارات ذاتية القيادة – بشكل كبير على التعلم العميق ومعالجة اللغات الطبيعية.

الذكاء الذكاء الاصطناعي أو ذكاء الآلة (MI) هو الذكاء الذي تعرضه الآلات، على النقيض من الذكاء الطبيعي (NI) الذي يعرضه البشر.

الذكاء الاصطناعي هي قدرة الكمبيوتر الرقمي أو الروبوت الذي يسيطر عليه الكمبيوتر على أداء المهام المرتبطة عادة بالبشر، إنها تقنية لديها القدرة على التفكير وحل المشكلات.

يركز الذكاء الاصطناعي بشكل أساسي على فهم وأداء المهام الذكية مثل التفكير وتعلم مهارات جديدة وتبني مواقف ومشاكل جديدة، وهو مزيج من علوم الكمبيوتر وعلم النفس والفلسفة.

يشير الذكاء الاصطناعي إلى العلوم والهندسة التي تستكشف كيفية محاكاة مختلف القضايا والوظائف في مجال الذكاء البشري، تغطي مجالات التكنولوجيا الذكاء الاصطناعي الإدراك والاعتراف والتفكير وعملية التعلم واللغة الطبيعية والترجمة الآلية والألعاب والشطرنج وما إلى ذلك.

من SIRI من Apple إلى السيارات ذاتية القيادة، الذكاء الاصطناعي يتقدم بسرعة، في حين أن الخيال العلمي غالبا ما يصور الذكاء الاصطناعي على أنها روبوتات ذات خصائص تشبه الإنسان، الذكاء الاصطناعي يمكن أن تشمل أي شيء من خوارزميات بحث Google إلى Watson من IBM إلى الأسلحة المستقلة، الذكاء الاصطناعي هو جزء رئيسي من العديد من التقنيات المتطورة، بما في ذلك الروبوتات والسيارات بدون سائق وعمليات البحث على الويب وألعاب الفيديو، تستخدم الذكاء الاصطناعي التقنيات خوارزميات متطورة، أو مجموعات من التعليمات، لحل المهام الصعبة للغاية .

تعمل بعض تقنيات الذكاء الاصطناعي خلف الكواليس لمعرفة من وماذا يحب الناس أثناء استخدامهم لوسائل التواصل الاجتماعي أو التسوق عبر الإنترنت.

يمكن تصنيف الذكاء الاصطناعي، الذي يعرف بأنه "قدرة النظام على تفسير البيانات الخارجية بشكل صحيح، والتعلم من هذه البيانات، واستخدام تلك الدروس المستفادة لتحقيق أهداف ومهام محددة من خلال التكيف المرن"(Kaplan 2022)، إلى ثلاثة أنواع: تحليلية، مستوحاة من الإنسان، وإنسانية، الذكاء الاصطناعي التحليلي له خصائص تتفق مع الذكاء المعرفي فقط. يحتوي الذكاء الاصطناعي المستوحى من الإنسان على عناصر الذكاء المعرفي والعاطفي، أي فهم العواطف البشرية، بالإضافة إلى العناصر المعرفية، ويستخدمها في

صنع القرار، الذكاء الاصطناعي إنساني يظهر خصائص جميع أنواع الكفاءات (أي الذكاء المعرفي والعاطفي والاجتماعي)، وهو قادر على أن يكون واعيا ذاتيا في تفاعلاته.

يمكن إرجاع ربيع الذكاء الاصطناعي، أي ولادته، إلى مقال آلان تورينج "آلات الحوسبة والذكاء" (1950) الذي وفر آلية لاختبار ذكاء الآلة، ومع ذلك، لم يتم صياغة مصطلح "الذكاء الاصطناعي" حتى عام 1956 في مشروع دارتموث الصيفي للبحوث حول الذكاء الاصطناعي (DSRPAI) الذي نظمه جون مكارثي ومارفن مينسكي، ما تلا ذلك كان بضعة الذكاء الاصطناعي "الصيف" و "الشتاء"، مجازا الصعود والهبوط اللاحق الذكاء الاصطناعي: في حين ساد التفاؤل العام والضجيج والتمويل الذكاء الاصطناعي المكثف خلال صيفها، كان العكس هو الحال خلال فصول الشتاء الذكاء الاصطناعي.

لم يصل خريف الذكاء الاصطناعي، أو فترة حصادها، إلا مؤخرا، عادة ما يعتبر "العام الأول" هو عام 2015، عندما تمكن نظام AlphaGo الذكاء الاصطناعي من Google من التغلب على البطل البشري في لعبة الطاولة المعقدة للغاية Go، كان يعتقد سابقا أن هذا الحدث مستحيل، وكان يمثل نقطة تحول في تاريخ الذكاء الاصطناعي، في السابق، لم تكن المعالجة الحسابية قوية بما يكفي للتعامل مع البيانات (الكبيرة) اللازمة لأنظمة الذكاء الاصطناعي قادرة على التعلم، يحتاج الذكاء الاصطناعي إلى الكثير من البيانات لتدريب الخوارزميات، وهنا يأتي دور وسائل التواصل الاجتماعي، لأنها تحتوي على كميات هائلة من البيانات.

2.1.2 التطبيقات الأساسية في مجال الذكاء الاصطناعي:

1.2.1.2 التعلم الآلي (ML)

يُمكِّن التعلم الآلي الآلات من "تعلم" مهمة ما من التجربة دون الحاجة إلى برمجتها على وجه التحديد حول هذه المهمة، (باختصار، تتعلم الآلات تلقائيًا دون إمساك اليد البشرية!) تبدأ هذه العملية بتزويدها ببيانات جيدة النوعية ثم تدريب الآلات عن طريق بناء نماذج مختلفة باستخدام خوارزميات مختلفة، يعتمد اختيار الخوارزميات على نوع المهمة التي نحاول أتمتتها.

ومع ذلك، بشكل عام، تنقسم خوارزميات التعلم الآلي إلى 3 أنواع، وهي: التعلم الخاضع للإشراف والتعلم غير الخاضع للإشراف والتعلم المعزز.

2.2.1.2 التعلم العميق

التعلم العميق هو مجموعة فرعية من التعلم الآلي، يتيح معالجة البيانات وإنشاء تنبؤات باستخدام الشبكات العصبية، ترتبط هذه الشبكات العصبية في شبكة / بنية تشبه الشبكات في الدماغ البشري.

تعني هذه البنية الشبيهة بالويب للشبكات العصبية الاصطناعية أنها قادرة على معالجة البيانات في نهج غير خطي، وهي ميزة مهمة على الخوارزميات التقليدية.

أحد الأمثلة على الشبكات العصبية العميقة هو RankBrain وهو أحد العوامل في خوارزمية بحث Google.

3.2.1.2 تعزيز التعلم

التعلم المعزز هو جزء من الذكاء الاصطناعي حيث تتعلم الآلة شيئًا بطريقة مشابهة لكيفية تعلم البشر، كمثال، افترض أن الآلة طالب، هنا يتعلم الطالب المفصل من أخطائه بمرور الوقت من خلال التجربة والخطأ.

هذا يعني أن الخوارزمية تقرر الإجراء التالي من خلال تعلم السلوكيات التي تستند إلى حالتها الحالية والتي ستزيد من المكافأة في المستقبل.

من الأمثلة الشهيرة على Reinforcement Learning برنامج Google alpha Go للكمبيوتر الذي تمكن من التغلب على بطل العالم في لعبة Go في عام 2017.

4.2.1.2 الروبوتات

الروبوتات هي مجال يتعامل مع إنشاء آلات تشبه البشر يمكنها أن تتصرف مثل البشر وتؤدي بعض الأعمال مثل البشر، الآن، يمكن للروبوتات أن تتصرف مثل البشر في مواقف معينة، لكن هل يمكنها أن تفكر مثل البشر أيضًا؟ هذا هو المكان الذي يأتي فيه الذكاء الاصطناعي قد تكون هذه الروبوتات قادرة على حل المشكلات في مجال محدود أو حتى التعلم في البيئات الخاضعة للرقابة.

مثال على ذلك هو Kismet، وهو روبوت للتفاعل الاجتماعي تم تطويره في مختبر الذكاء الاصطناعي التابع لمعهد ماساتشوستس للتكنولوجيا، يتعرف على لغة جسد الإنسان وأيضًا صوتنا ويتفاعل مع البشر وفقًا لذلك يمكنه محاكاة السعادة والحزن والدهشة والتغيرات في حركة الرأس (عبده، 2021، صفحة 243)، مثال آخر هو Robonaut، الذي طورته وكالة ناسا للعمل جنبًا إلى جنب مع رواد الفضاء في الفضاء.

5.2.1.2 معالجة اللغة الطبيعية (NLP)

من الواضح أنه يمكن للبشر التحدث مع بعضهم البعض باستخدام الكلام وأصبح هذا ممكنًا للآلات أيضًا! يُعرف هذا باسم معالجة اللغة الطبيعية حيث تقوم الآلات بتحليل وفهم اللغة والكلام كما يتم التحدث بهما (الآن إذا تحدثت إلى آلة فقد تتحدث مرة أخرى!)، هناك العديد من الأجزاء الفرعية من البرمجة اللغوية العصبية التي تتعامل مع اللغة مثل التعرف على الكلام وتوليد اللغة الطبيعية وترجمة اللغة الطبيعية وما إلى ذلك، فهي تتضمن تحليل وتفهم وخلق استجابات لتمكين التفاعل مع استخدام البشر النظم المتاحة لهم بدل لغة الحاسوب (الهادي، 2021، صفحة 317).

إن البرمجة اللغوية العصبية (NLP) تحظى بشعبية كبيرة في الوقت الحالي لتطبيقات دعم العملاء، ولا سيما chatbot، تستخدم روبوتات المحادثة هذه ML و NLP للتفاعل مع المستخدمين في شكل نصي وحل استفساراتهم، لذلك تحصل على اللمسة الإنسانية في تفاعلات دعم العملاء دون التفاعل المباشر مع أي إنسان، بعض الأمثلة الأكثر شيوعًا لتطبيقات البرمجة اللغوية العصبية هي Alexa من Amazon و Siri من Apple.

6.2.1.2 نظم التوصية

عندما تستخدم Netflix، هل تحصل على توصية بالأفلام والمسلسلات بناءً على اختياراتك السابقة أو الأنواع التي تحبها؟ يتم ذلك من خلال أنظمة التوصية التي توفر لك بعض الإرشادات حول ما يجب أن تختاره بعد ذلك من بين الخيارات الواسعة المتاحة عبر الإنترنت.

يمكن أن يعتمد نظام التوصية على توصية قائمة على المحتوى أو حتى التصفية التعاونية، تتم التوصية القائمة على المحتوى من خلال تحليل محتوى جميع العناصر، على سبيل المثال يمكنك التوصية بالكتب التي قد تعجبك بناءً على وصف العنصر وملفك الشخصي الأساسي من ناحية أخرى، تتم عملية التصفية التعاونية عن طريق تحليل سلوك القراءة السابقة لأشخاص مثلك ثم التوصية بالكتب بناءً على ذلك.

7.2.1.2 رؤية الحاسوب

لم يكن التقاط صورة ومشاركتها أسهل من أي وقت مضى، في الواقع يتم تحميل ملايين الصور وعرضها يوميًا على الإنترنت، لتحقيق أقصى استفادة من هذا الكم الهائل من الصور عبر الإنترنت، من المهم أن تتمكن أجهزة الكمبيوتر من رؤية الصور وفهمها، وبينما يمكن للبشر القيام بذلك بسهولة دون تفكير، فإنه ليس بهذه السهولة على أجهزة الكمبيوتر! هذا هو المكان الذي يأتي فيه Computer Vision.

تستخدم رؤية الكمبيوتر الذكاء الاصطناعي لاستخراج المعلومات من الصور، يمكن أن تكون هذه المعلومات عبارة عن اكتشاف كائن في الصورة، وتحديد محتوى الصورة لتجميع الصور المختلفة معًا، أحد تطبيقات رؤية الكمبيوتر هو التنقل للمركبات ذاتية القيادة من خلال تحليل صور المناطق المحيطة مثل AutoNav المستخدمة في سبيريت وسبورتونيتي المتجولتين اللتين هبطتا على سطح المريخ.

8.2.1.2 إنترنت الأشياء

يتعامل الذكاء الاصطناعي مع إنشاء أنظمة يمكنها تعلم محاكاة المهام البشرية باستخدام خبرتهم السابقة ودون أي تدخل يدوي، إنترنت الأشياء، من ناحية أخرى، عبارة عن شبكة من الأجهزة المختلفة المتصلة عبر الإنترنت ويمكنها جمع البيانات وتبادلها مع بعضها البعض، الآن، تولد جميع أجهزة IoT هذه الكثير من البيانات التي يجب جمعها واستخراجها للحصول على نتائج قابلة للتنفيذ (أبو غزالة، 2019)، هذا هو المكان الذي يأتي

فيه الذكاء الاصطناعي إلى الصورة، يتم استخدام إنترنت الأشياء لجمع ومعالجة كمية هائلة من البيانات التي تتطلبها خوارزميات الذكاء الاصطناعي، في المقابل تحول هذه الخوارزميات البيانات إلى نتائج مفيدة قابلة للتنفيذ يمكن تنفيذها بواسطة أجهزة IoT.

2.2 تحليل المحتوى الرقمي العربي:

كانت ممارسة تحليل المحتوى في الممارسة العملية منذ أوائل الأربعينيات، بعد ذلك، كان يقتصر على فحص نصوص الدراسات المختلفة يدويًا أو من خلال أجهزة الكمبيوتر المركزية البطيئة للتحقق من تكرار المصطلحات / الكلمات المحددة، في وقت لاحق، في الخمسينيات من القرن الماضي، تضمن تحليل المحتوى وسائل معقدة للتحليل تستخدم للتركيز على المفاهيم والعلاقات الدلالية أيضًا. في الوقت الحاضر، مع ثورة التكنولوجيا، يتم استخدام تحليل المحتوى لتحليل جوانب مختلفة من المحتوى لاستكشاف النماذج العقلية جنبًا إلى جنب مع الأهمية المعرفية واللغوية والثقافية والاجتماعية والتاريخية للمحتوى.

تحليل المحتوى عبر الإنترنت أو التحليل النصي عبر الإنترنت إلى مجموعة من تقنيات البحث المستخدمة لوصف وتقديم استنتاجات حول المواد المتاحة عبر الإنترنت من خلال الترميز والتفسير المنهجي، تحليل المحتوى عبر الإنترنت هو شكل من أشكال تحليل الاتصالات القائمة على الإنترنت.

1.2.2 الجوانب المفاهيمية لتحليل المحتوى:

1.1.2.2 ماهية تحليل المحتوى:

تحليل المحتوى هو أسلوب كمي ونوعي يقدم تقييمًا أكثر موضوعية للمحتوى، سيكون بالتأكيد أكثر دقة من المقارنة المبنية على انطباعات أي مستمع، إنه أكثر فعالية من المراجعة أو التقييم، ستجد الأرقام والنسب الأساسية لقياس أداء المحتوى الخاص بك.

يمكن استخدام تحليل المحتوى من أجل – إزالة الموضوعات من الملخصات الخاصة بك وتبسيط اكتشاف الاتجاهات في مكانتك.

لتوجيه تحليل المحتوى، تقوم عمومًا بجمع البيانات من النصوص الشفوية أو المرئية أو المكتوبة، يوجد هذا النوع من البيانات في الكتب والمجلات والصحف والخطب والأفلام والمقابلات والصور الفوتوغرافية ووسائل التواصل الاجتماعي والويب وأنواع أخرى مختلفة من المحتوى.

بشكل عام، يعد تحليل المحتوى أسلوبًا خبيرًا يساعد في معرفة الغرض من أي شكل من أشكال محتوى الاتصال و تأثيراته ورسائله.

دعونا الآن نلقي نظرة على أنواع النصوص في تحليل المحتوى–

- النص المكتوب على سبيل المثال الأوراق والكتب

- نص شفوي للأداء المسرحي الفوري والكلام
- نص أيقوني مثل اللوحات والرسومات والأيقونات
- نصوص صوتية ومرئية مثل البرامج التلفزيونية ومقاطع الفيديو والأفلام
- النصوص التشعبية مثل النصوص الموجودة على الإنترنت

2.1.2.2 الأسئلة التي يجب أن يجيب عليها كل تحليل محتوى

وفقًا لاقتراح (Krippendorff, 2018)، يجب أن تكون هناك ستة أسئلة يجب معالجتها من خلال كل تحليل للمحتوى–

- ما هي البيانات أو المعلومات التي يتم التحقيق فيها أو تحليلها؟
- كيف يتم تعريف أو تمييز البيانات والمعلومات ذات الصلة؟
- من أي تعداد أو أنواع من السكان يتم استخلاص البيانات؟
- ما هو السياق المناسب للمحتوى؟
- ما هي حدود تحليل المحتوى الخاص بك؟
- ما الذي يجب قياسه بتحليل المحتوى؟

3.1.2.2 أنواع تحليل المحتوى

كما تمت مناقشته أعلاه، فإن الكمية والنوعية هما شكلين من أشكال تحليل المحتوى، وسيساعدك الاختلاف بينهما على فهم أهمية هذين التحليلين–

1.3.1.2.2 تحليل المحتوى الكمي

يمكنك استخدامه من خلال التركيز على العد وقياس حدوث عبارات وكلمات ومفاهيم وموضوعات محددة، على سبيل المثال، إذا كنت تقوم بتحليل المحتوى لخطاب حول قضايا التوظيف، فسيتم تركيز وتحليل مصطلحات مثل الوظائف والبطالة والعمل وما إلى ذلك، فهو تقنية بحث للوصف الكمي الواقعي والمنهجي بالنسبة للمحتوى التواصلي الظاهري (غاريسون، اندرسون، و الأبرش، 2006).

2.3.1.2.2 تحليل المحتوى النوعي

يركز هذا النوع من تحليل المحتوى على تفسير وفهم نوع معين من المحتوى، على سبيل المثال، إذا أجرينا تحليلًا نوعيًا على مثال خطاب مشكلة التوظيف المذكور أعلاه، فستبحث عن مصطلح البطالة ومصطلحات أخرى (عدم المساواة، الاقتصاد، إلخ) بجواره، بعد ذلك، يجب عليك تحليل علاقات هذه المصطلحات لقياس النوايا والعلاقات الدلالية لهذه المصطلحات والمفاهيم في الحملات، يمكن فهم هذين النوعين من تحليل المحتوى على أنهما التحليل المفاهيمي والتحليل العلائقي.

أولا. التحليل المفاهيمي

يشبه التحليل المفاهيمي التحليل الكمي ويتم إجراؤه بطريقة محددة، أثناء القيام بتحليل المحتوى المفاهيمي، يتم اختيار مفهوم للفحص، وتشير الدراسة إلى تحديد مدى وجوده وتسجيله.

- **مثال على تحليل المحتوى المفاهيمي**

على سبيل المثال، لنفترض أن لديك انطباعًا بأن مؤلفك المفضل كثيرًا ما يكتب عن الحب، لذلك من خلال التحليل المفاهيمي، يمكنك تحديد عدد المرات التي تظهر فيها كلمات مثل الإعجاب والولع والإعجاب والعشق في النص.

ثانيًا. تحليل العلاقات

بينما في التحليل العلائقي، يبدأ بتحديد الأفكار الموجودة بالفعل في النص المحدد أو مجموعة المستندات، إنه مشابه تمامًا للتحليل النوعي، يتعامل مع فحص العلاقات بين المفاهيم والمصطلحات في المحتوى.

- **مثال على تحليل المحتوى العلائقي**

بالعودة إلى مثال "الحب" نفسه، تبدأ بالخطوة الأولى وتفحص علاقة المحتوى، يمكنك تحديد هذه الكلمات (مثل الإعجاب، والولع، والإعجاب، والعشق) ثم تستنتج ما هي المعاني المختلفة التي تظهر من هذه المجموعة من الكلمات، عندها تكمل أن يكتب مؤلفك المفضل عن الحب في كثير من الأحيان.

هذا ويمكننا أن نقول أن التحليل المفاهيمي يركز على النظر في حدوث المصطلحات المختارة في النص، على الرغم من أن الوقت يمكن أن يكون ضمنيًا أو صريحًا أيضًا، من ناحية أخرى، يركز التحليل العلائقي على البحث عن العلاقات الدلالية أو الأساسية، يُنظر إلى المفاهيم الفريدة، في حد ذاتها، على أنها ليس لها أهمية متأصلة، بدلاً من ذلك، فإن الضمني هو نتاج العلاقات بين المفاهيم في النص.

4.1.2.2 استخدامات تحليل المحتوى

يستخدم الباحثون تحليل المحتوى للتعرف على رسائل وأغراض وتأثيرات محتوى الاتصال، ربما يمكن تطبيقه لفحص أي جزء من الكتابة أو حدوث محادثة مسجلة، لذلك هذه هي الحقيقة المركزية، يستخدم الباحثون تحليل المحتوى، يمكنهم أيضًا وضع افتراضات حول المنتج وجمهور النصوص لتحليلها، علاوة على ذلك يتم استخدام تحليل المحتوى في مجموعة متنوعة من المجالات، تتراوح من قضايا الجندرة والعمر، إلى علم النفس والعلوم العقلية، ودراسات التسويق والإعلام، والأدب والبلاغة، وعلم الاجتماع والعلوم السياسية، والإثنوغرافيا والدراسات الثقافية، والعديد من المجالات الأخرى من سؤال، أيضًا، يشير تحليل المحتوى إلى وجود علاقة وثيقة مع علم اللغة الاجتماعي والنفسي ويلعب دورًا أساسيًا للغاية في تحسين الذكاء الاصطناعي.

أولاً. يمكنك استخدام تحليل المحتوى لعمل استنتاجات بشأن أسلاف الاتصال مثل:

- تحليل سمات الأفراد
- استنتاج الجوانب الثقافية والتغيير
- تقديم الأدلة القانونية والتقييمية
- الإجابة على أسئلة التأليف المتنازع عليها

ثانيًا. يتم استخدام تحليل المحتوى أيضًا لوصف وتقديم الاستنتاجات المرتبطة بخصائص أي اتصال مثل:

- وصف الاتجاهات في محتوى الاتصال
- ربط الخصائص المعروفة للمصادر بالرسائل التي تنتجها
- مقارنة محتوى الاتصال بالمعايير
- إثبات علاقة الخصائص المعروفة للجماهير بالرسائل المنتجة لهم
- التعبير عن أنماط الاتصال المختلفة
- تقويم أساليب الإقناع

يمكن أيضًا استخدام تحليل المحتوى لإجراء استنتاجات بشأن تأثيرات وعواقب الاتصال مثل–

- قياس سهولة القراءة
- تحليل تدفق المعلومات
- تقييم الردود على الاتصالات

2.2.2 جوانب من تحليل المحتوى العربي في ظل تقنيات الذكاء الإصطناعي

بين أفضل 30 لغة يتم التحدث بها في أنحاء العالم تحتل اللغة العربية المرتبة الخامسة، كما تحتوي اللغة العربية على ثلاثة تنويعات، والتي تشمل اللغة العربية الفصحى، واللغة العربية الفصحى الحديثة (MSA)، واللغة العربية العامية. عادة ما تستخدم اللغة العربية الفصحى للنصوص الدينية والتاريخية. يمكن العثور على MSA في النص العربي المكتوب اليوم في معظم القنوات الرسمية. اللغة العربية العامية أو الجدلية هي اللغة المنطوقة في قنوات التواصل الاجتماعي وغير الرسمية. بالإضافة إلى ذلك، تختلف اللهجات من بلد عربي إلى آخر.

مع الزيادة السريعة في حجم المعلومات على الويب، جذب تصنيف النصوص انتباه العديد من الباحثين لاستخدام تصنيف النص (التصنيف) كوسيلة لتبسيط الوصول إلى المعلومات المفيدة. يمكن تعريف تصنيف النص كإحدى مهام التنقيب عن النص الرئيسية على أنه عملية تعيين فئة (تسمية) محددة مسبقا إلى مستند (نص) غير مسمى استنادا إلى محتواه. تم استخدام تصنيف النصوص للعديد من التطبيقات مثل تحسين أداء أنظمة استرجاع المعلومات، وتصفية الرسائل غير المرغوب فيها، وفي أنظمة المعلومات الطبية.

ومن الناحية العملية، يتضمن نظام تصنيف النصوص النموذجي أربع مراحل رئيسية، حيث يمكن أن تشمل كل مرحلة عدة خطوات أخرى. وتشمل هذه المراحل مرحلة المعالجة المسبقة للنص، والتي تتضمن عدة خطوات تهدف إلى إعداد النص ليكون جاهزا للمعالجة بواسطة نموذج التصنيف، عادة تتضمن هذه المرحلة بعض تقنيات معالجة اللغات الطبيعية الكلاسيكية الأخرى مثل التحليل، والترميز، وإيقاف إزالة الكلمات، وترجيح المصطلحات، والسيقان، وجزء من وضع علامات الكلام.

ترميز النص هو عملية تقسيم سلسلة من السلاسل إلى أجزاء مثل الكلمات والكلمات الرئيسية والعبارات والرموز والعناصر الأخرى التي تسمى الرموز المميزة. نتيجة عملية الترميز هي الرموز المميزة أو المصطلحات التي تجعل من السهل حساب عدد المصطلحات نفسها في كل مستند أو أي عملية حسابية أخرى مطلوبة لحساب وزن كل مصطلح في كل مستند، إيقاف إزالة الكلمات هو عملية إزالة علامات الترقيم وعلامات التنسيق والأرقام وحروف الجر والضمائر والعطف والأفعال المساعدة. يمكن تعريف كلمات الإيقاف بأنها مجموعة من الكلمات التي تحتوي على تردد عال في مجموعة المستندات ولن تساعدنا في عملية التصنيف. تتميز إزالة كلمات الإيقاف بالعديد من المزايا، مثل تقليل حجم مجموعة المستندات والسماح بتحديد المصطلح الأكثر شيوعا في كل مستند. تتضمن اللغة العربية عدة أنواع من كلمات التوقف، إلى جانب ذلك، تعتبر الكلمات العامة والأرقام والرموز والأحرف الخاصة أيضا كلمات توقف في العديد من التطبيقات.

يجب أن تتعامل تطبيقات اللغة العربية الطبيعية واستخراج النصوص مع العديد من المشكلات المعقدة ذات الصلة بطبيعة اللغة العربية وبنيتها. على سبيل المثال، عملية الترميز للنص العربي ليست مهمة مباشرة لأن اللغة غنية من الناحية المورفولوجية والكلمات مضغوطة، حيث يمكن أن تتوافق الكلمة مع عبارة أو جملة بأكملها.

لهذه الأسباب، تحتاج اللغة العربية إلى معالجة مسبقة دقيقة لأنها تحتوي على بعض الميزات التي تختلف عن اللغات الأخرى. إلى جانب ذلك، قد تؤثر هذه التحديات على نتائج أي عملية تحليل نص مثل التصنيف أو تحليل المشاعر (Duwairi & El-Orfali, 2014, p. 509).

قد يحتوي النص العربي أيضا على كلمات أو حروف أو جمل من لغة أخرى مثل اللغة الإنجليزية. قامت معظم التطبيقات بإزالتها من النص قبل المعالجة.

يتم تعريف Stemming على أنه العملية التي ترجع جزء الكلمة المتبقي بعد إزالة بعض البادئات واللواحق من الكلمة. تستخدم هذه العملية عادة لتقليل حجم مجموعة المصطلحات لكل وثيقة في مجموعة وثائق التدريب المستخدمة في عملية التصنيف.

3 تحديات الذكاء الإصطناعي في تحليل المحتوى الرقمي العربي في البحوث الاجتماعية

1.3 تحديات منهجية

1.1.3 عدم توافق البيانات المستخدمة في أنظمة الذكاء الاصطناعي مع الواقع

بين جمع البيانات والمخرجات المتأتّية من البرامج الذكية(Outputs) ، هناك العديد من عمليات الترجمة التي يتم فيها تجميع البيانات ومعالجتها وإجراء العمليات الحسابية وعرض النتائج بصريًا، وفي سياق عمليات الترجمة المعقدة هذه، يتم إنشاء البنية التكنولوجية للواقع والتي يمكن أن تتناقض مع واقع الأفراد أو المجموعات.

إنّ جمع البيانات وحده ينطوي دائمًا على الاختيار بين العوامل التي يتمّ قياسها، وتلك التي يجري تجاهلها، (Gitelman, 2013, p. 10) وبالتالي لا يتمّ استخدام البيانات الشخصية إلّا لبناء الخصائص والاهتمامات وغيرها، لدى الشخص المعني، ويتمّ تقسيم الأفراد بناءً على تدفّقات بيانات مختلفة، وبعد ذلك وباستخدام تصنيفات وفئات مختلفة تتم إعادة بناء "شخصية رقمية"، أي تكوين ملف تعريف افتراضي لشخص معين (David, 2020, p. 1)، وبهذه الطريقة يكون للبيانات تأثير فاعل (Gitelman, 2013)، لكن هذا يحمل في طياته ثغرات، ذلك أنه ينطوي على خطر أن تؤدي عمليات جمع ومعالجة البيانات إلى بناء شخصية افتراضية "مجزّأة" أو "مشوّهة" أو "غير صحيحة" من جوانب معينة، ما قد يتناقض بشكلٍ واضح مع صورة هوية الشخص الحقيقي (Maria, 2006, p. 68)، خصوصًا أنّ البيانات الشخصية عادةً تكون ناتجة من سياق منصات التواصل الاجتماعي، التي غالباً ما تتّسم بتمثيلٍ غير دقيق أو ضعيف نسبيًا (danah & Kate, 2012, p. 667)، كما يمكن أن تكون مجموعات البيانات متحيّزة، حيث قد يتمّ تزييف البيانات بسبب أخطاء الجهاز أو بواسطة حسابات وهمية أو برامج روبوت إلخ،.. والنتيجة في هذه الحالة ستكون بيانات مجتزأة ومشوّهة يتم جمعها ومعالجتها ودمجها في أنظمة التعلّم الآلي، وعليه، فإنّ أحد التحديات التي تواجه أنظمة الذكاء الاصطناعي هي النتائج المنبثقة عن هذه البيانات ما يمكن أن يتناقض تمامًا مع معطيات الواقع، وهذا شيء نادر ما يفكر فيه الناس.

2.3 تحديات إجتماعية

1.2.3 المعرفة المحدودة لدى مطوري التطبيقات

يعمل مهندسي الذكاء الاصطناعي على تطوير العديد من التطبيقات في شتى المجالات، والتي لها تأثير على عددٍ كبير من الأشخاص في كثيرٍ من الأحيان، إلا أن غالبية هؤلاء المهندسون لا يملكون عادةً إلّا خلفيةٍ علمية في علوم الحاسوب فقط (Emanuelle, et al., 2017, p. 30)، وعلم محدود في باقي العلوم، ويمكن أن يصادف مهندسي البرمجيات حالات اجتماعية أو نفسية أو أخلاقية أو حتى سياسية تكون لديهم خلفية غير كافية، الشي الذي يخلق صعوبات –نقاطٍ عمياء أو سوء استيعاب لحالات معينة– في فهم التفاعلات

والسياقات ضمن المجالات التي سيتمّ فيها تطبيق التكنولوجيا، لذلك، بل حتى في مرحلة تجريب أو تطبيق التقنيات الرقمية المختارة، تقع أحيانًا أشياء غير مرغوب بها أو مخطّط لها (Hyysalo, 2006, p. 621)، لهذا يحتاج مهندسي الذكاء الإصطناعي إلى فهمٍ منهجي للمجتمعات، والعلاقات الإنسانية، وكذلك الأخلاقية، طبعاً لا ننفي أنّ المجتمعات العلمية التقنية تستحقّ التقدير، مع الاعتراف بدور تقاطع علم الحاسوب مع شتى العلوم.

2.2.3 تصدير القيم الضمنية لتطبيقات الذكاء الاصطناعي

لدى مهندسي ومطوري تطبيقات الذكاء الاصطناعي قيم خاصة اجتماعية وغيرها تنعكس رغم تجردهم على الأفكار ضمنيا في شكل أو الأثر التي تتركه التطبيقات في حياة الناس ممن يطلبون على خدماتهم، فظروف تنشئتهم الاجتماعية وثقافاتهم تنتقل بشكل متزايد على وظائف تلكم التقنيات بل وتخلق اعتبارات وتساؤلات حول مثلا الشفافية والتحيز والعدالة والإنصاف، وهو ما يجعل بناء تطبيقات عادلة وغير تميزية أمر صعب المنال، وهنا حقيقة تتظهر جليا للعيان، بأن التطبيقات ليست المسؤولة في حد ذاتها على هذه التبعات بل المهندسين والمطورين، فهم من يعمل على بناء المعادلات ولغة البرامج والتطبيقات وتكويدها (Software Coding)، وحيز العمل والبيئة التي يتم بناء وابتكار هذه التطبيقات فيها لثقافتها دور وأثر سواء في مرحلة التدريب أو التجريب أو التنفيذ، يجري تضمين كل هذه القيم ضمن التطبيق النهائي (Kitchin, 2017, p. 25).

3.2.3 محدودية التنوع في جنس ولون العاملين في بحوث وصناعات الذكاء الإصطناعي:

تطغى على نوع العلمين في المجال التقني، من أساتذة لمختلف جامعات العالم الرائدة مثل "ستانفورد" أو "أكسفورد" أو "بيركلي" فما نسبته 80% من مجموع هؤلاء العاملين لديهم نوع الجنس ذكر، على مستوى الولايات المتحدة مثلاً، يشكّل الرجال أكثر من 70% من مجموع المتقدمين للحصول على وظائف الذكاء الاصطناعي، هناك أيظًا 14% فقط من الموظفين من ذوي الأصول الأسبانية أو ذوي البشرة الداكنة رغم أن نسبتهم 27% من إجمالي عدد السكان(Khan, Gawalt, & CooK, 2016) ، وتأكد هذا التفاوت أيظا المعهد الإحصائي بألمانيا، حيث تصل نسبة النساء بين طلبة علوم الحاسوب حوالي 20% فقط ضمن السنة الأولى، ما يستدعي زيادة ملحة في نسبة مطوري التطبيقات من الإناث، من أجل تمثيل أفكارهن وقيمهنّ بشكلٍ أفضل في مجتمعات الذكاء الاصطناعي (Lundberg & Stearns, 2018, p. 11).

4.2.3 محدودية عدد الخبراء والكفاءات في الذكاء الاصطناعي

الكفاءات من يستطيعون البحث وتطوير برامج الذكاء الاصطناعي هم قلة قليلة حول العالم، ذلك أنهم يستعملون عدة برامج مع بيانات محددة ضمن سياق خاص فهذا التكيف يلزم خبراء وتقنيين ذوي مهارات فائقة خاصة في مجال التعلم الآلي والذكاء الاصطناعي (Markow, Hughes, & Miller, 2017)، فهم نادرون حول العالم وهذا هو تفسير تهافت الشركات لتوظيفهم.

5.2.3 الرفض المجتمعي لبعض التطبيقات الذكية:

رغم وصول الذكاء الاصطناعي لابتكار العديد من الروبوتات التي يمكنها أن تقوم بمهام كالطبيب في اكتشاف الامراض إلا أن نسبة غير عادية من المرضى يفضلون الطبيب البشري في القيام بعملية التشخيص رغم حدوث بعض الأخطاء لديهم في مقابل محدودية أو انعدام ذلك لدى الروبوت الذكي مثل روبوت "دافنشي" الجراحي، أو الروبوتات التي خضعت للتعلم الآلي عن طريق البيانات الشخصية لافراد متوفين تم تزويدها لتجسيدات رقمية متحركة ثلاثية الأبعاد، لتقوم بالمحادثة مع افراد العائلة او التذكير بمهام ذاخل البيت بعد وفاتهم مثل تطبيق "Replika" (Metz, 2017) .

6.2.3 التكاليف العالية والخفية للتطبيقات والأنظمة الذكية

يضطر أحيانا شركات الذكاء الاصطناعي زيادة

عن كونها تعمل على عمليات تقنية، بل يعمل تتعدى فكرى الذكاء الاصطناعي الى فرض شخصية "أسطورية" متكاملة، ومصدر هذه النظرة هو زيادة الأتمتة للإجراءات والعمليات الأكثر تعقيدًا، ومع ذلك، لا ينحصر عمل الأنظمة الذكية في الواقع بعمل مطوري التكنولوجيا الذين يكتسبون أجوراً عاليةً في شركات التكنولوجيا والجامعات بل يرتبط أيضًا بعدد من التكاليف الخفية غير الواضحة، فمثلًا لا بدّ من جمع البيانات التي تعد الركيزة الأساسية، وهذا يحتاج إلى مجموعات عمل لا تربطها بشركات الذكاء الصناعي علاقة مباشرة، بل تتحصل عليها من دول نامية على غرار بنغلادش ونيبال بل ومن الصين والهند، وهناك تكاليف مستترة أخرى، إذ تتطلّب الأنظمة توافر أجهزةٍ بمواصفات خاصة لها القدرة على تشغيلها، وعليه لابد من صناعة هذه الأجهزة بطريقة ما، من مواد أولية وتزويدها بالطاقة .(CASILL, 2017)

7.2.3 منظومة وبنى تحتية متقلبة ومتطلّبات مادية عالية

تتطلب اليوم أنترنت الأشياء التدعيم ببرامج ذكية، فقد انخرطت إشارات الطرق وأنظمة الأسلحة والأقمار الصناعية والمستشفيات والسيارات والطائرات والسفن ومحطات الطاقة والآلات الصناعية والزراعية إلى شبكة انترنت الأشياء، فهذا يجعل من عامل السرية والنزاهة المطلوبة حيال البيانات التي يتم معالجتها غير مضمونة الحدوث، بل غالبًا ما تعرض هذه الأنظمة للهجمات الإلكترونية، ما يحتم إنشاء أنظمة أكثر تعقيدًا وأمانًا من جهة، وانسابية وتخفيف التعقيد في الجهة المقابلة، ولتحقيق ذلك، لابد من توفير عدد من المتطلّبات المادية، ومن الواضح أنّ المتطلبات الأساسية لتحقيق نظم ذكية تكلفتها عالية جدًا، فهناك حاجة لاستقدام المواد الخام من أجل إنتاج أجهزة حاسوبية، ومطلوب أيضاً إمدادات مستمرة من الطاقة لعددٍ لا يحصى من الحواسيب، وكذلك توافر شبكات إلكترونية تربط بين هذه الحواسيب؛ ولا بدّ أن يكون هناك أساس لكمياتٍ كبيرة من البيانات التي يمكن الاستعانة بها لتدريب الخوارزميات، والتي يفترض بدورها تشغيل الملايين من أجهزة الاستشعار (المجسّات)

من أجل انترنت الأشياء، وبالتالي، هناك عدد من الشروط الأساسية فيما يتعلق بالبنية التحتية المستقرة بحيث يمكن تطوير وتشغيل أنظمة الذكاء الاصطناعي بالشكل المطلوب، ونظرًا للتطورات السياسية والإيكولوجية الحالية حول العالم، ليس هناك ما يضمن أنّ هذه البنية التحتية ستحافظ على الاستقرار في المستقبل .

3.3 تحديات تقنية

1.3.3 التفكير البشري مختلف تمامًا عن الآلات الذكية

رغم أنه تجري المقاربة بين الدماغ والآلة، إلّا أنّ الشبكات العصبية الحالية التي يمكن أن تكون الأكثر تطوراً في بناء الآلات الذكية لا تمثّل سوى شيء بسيط جدًا من أنسجة المخ البشري، وإذا كان الدماغ البيولوجي هو الهيكل الذي يتمّ من خلاله تعريف الذكاء، فإن الذكاء الاصطناعي لا يزال في بداياته.

وعند الكتابة عن الذكاء الاصطناعي، كثيرًا ما يُقال إن الأجهزة الحاسوبية عبارة عن "آلات تفكير" قادرة على "فهم الأشياء وتعلّمها"، وفي كثيرٍ من الأحيان، يتمّ رسم القياس وفق الدماغ البشري، أو يقف المخ كرمزٍ للذكاء الاصطناعي، لكن في الواقع، تختلف الأساليب الإحصائية للتعلّم الآلي بطرقٍ عديدة عن المفاهيم البيولوجية للتفكير أو الفهم أو التعلّم، وفي سبعينيات القرن العشرين اعترض الفيلسوف الأميركي "هوبرت دريفاك" في كتابه الكلاسيكي"What Computers Can't Do" ؟" أي (ما الذي لا تستطيع أجهزة الحاسوب فعله) في سبعينيات القرن الماضي، على افتراضٍ مفاده أنّ الدماغ ممكن معالجته وفقًا لقواعد، ومن ثمّ يمكن إعادة استنساخه عن طريق محاكاة الحاسوب، وقد لاقت اعتباراته صدًى حتى يومنا هذا، لذلك، فعند الحديث عن برامج "التعرّف على الوجوه" على سبيل المثال، سيكون من الأدقّ القول إنها "عملية ضبط لمجموعة من قيم البكسل (Pixels) المرتبطة في كثيرٍ من الأحيان بوجود الوجوه في بيانات التدريب التي تم جمعها سابقاً، ذلك لأنّ أجهزة الحاسوب لا تتعرّف أو تدرك أو تعلم أساساً ما هو الوجه، ومن المهم أن نذكر أنّ هناك مقترحات بشأن كيفية تحقيق تقارب أكبر بين التفكير والتعلّم بين البشر والآلة، وقد اقترح باحثون ثلاث خطوات لنهج يشمل كلا العنصرين:

أولاً: لا بدّ من تحويل التركيز بعيدًا عن تطوير أساليب التعرف على الأنماط، والاتجاه نحو التركيز على العلاقات السببية.

ثانياً: يجب العمل على تعليم الآلات السمات الأساسية للنظريات الفيزيائية والنفسية حتى يتسنّى إثراء المعرفة المولّدة بخلفية معرفية مناسبة، وبنفس الطريقة التي لا يدرك بها الناس ببساطة البيانات الحسيّة البسيطة، ولكنهم يفسرونها دائمًا، ولذلك ينبغي أن لا تقوم الأجهزة الذكية بمعالجة أجزاء البيانات المجرّدة فحسب بل أيضًا الكشف عن السياقات الشاملة للمعنى.

ثالثاً: ينبغي للآلات تعلّم كيف تتعلّم بحيث يمكن تطبيق المعرفة بشكلٍ أسرع وأكثر دقة على المواقف والمهام الجديدة.

فضلًا عن ذلك، هناك بعض الاختلافات الأخرى بين القدرات العقلية للإنسان والقدرات "المعرفية" لأجهزة الحاسوب، ذلك أنّه رغم التطورات السريعة في مجال الذكاء الاصطناعي، يتفوق البشر حتى الآن على هذه اللأجهزة، في مهام مثل تعلّم اللغة، الإبداع، الحدس، التعرّف البصري على السياقات المعقدة وغير ذلك، وفي حين أنّ الناس يتشابهون في عمليات التعلم بفهمٍ أساسي لمفاهيم عامة مثل الفضاء، أو الوقت، أو العدد، أو ديمومة الأشياء أو السببية... فإنّ أجهزة الحواسيب لا تمتلك هذه المعرفة الأساسية ويجب تنفيذها ميكانيكيًّا من خلال البرمجة المناسبة.

2.3.3 العديد من خوارزميات التعلّم لا تتّسم بالمرونة في وظائفها

قبل استخدام التطبيقات الذكية، لا بدّ أولًا من تدريب الخوارزميات الأساسية، ومع ذلك، تؤدي عملية التدريب مع مجموعات البيانات الممثّلة إلى تخصّص خوارزميات التعلّم تلك، وهذا يعني أنه بمجرد تغيير اللون، أو تغيير قاعدة في لعبة ما، أو حذف الأحرف في النصوص سيؤدي إلى معالجة غير صحيحة للآلة38، ولكن بالنسبة للبشر فإنّ التكيّف مع مثل هذه التغييرات سيكون سهلاً للغاية، وصحيحٌ، أنه يمكن للبرمجيات التي يتمّ تدريبها، عبر اعتماد التكرار الكافي للعبة من تطوير قدراتها وهزيمة خصومها المحترفين من البشر، لكن تطبيقات التعلّم الآلي ما تزال محصورة في تلقّي مدخلات محددة للغاية، كما أنها تحتاج إلى تحقيق أهدافٍ تكون محددة أيضًا، وبينما يستطيع اللاعبون من البشر التكيّف بمرونة مع الاختلافات في اللعبة، أو دخول قواعد جديدة وذلك في غضون فترةٍ زمنية قصيرة للغاية، تحتاج أجهزة الحاسوب إلى مراحل التدريب الخاصة بها، وهذا ما يكون نتيجته إنفاق وقت وموارد مختلفة، وحقيقة أن اللاعبين البشريين هم أكثر مرونةً، فيمكن تفسيرها أنه لديهم بعض المعرفة السابقة الواقعية عن العلاقات السببية، والهندسة وما إلى ذلك، في حين أنّ خوارزمية التعلّم في كثير من الحالات تبدأ عملية التعلّم من الصفر.

3.3.3 قيود تصنيف البيانات على أنظمة الذكاء الاصطناعي

إنّ أحد المجالات التي يتم فيها الإشادة بإمكانيات أنظمة الذكاء الاصطناعي على وجه الخصوص هو مجال التعرّف على الصور، وبناءً على المراجعات الأدبية، فالأبحاث التي توضح كيف تتفوق أدوات التعرّف على الصور على المشغّلين من البشر، كما الحال في مجال الأشعة مثلًا، تولّد باستمرار الانطباع بوجود قدرات بصرية "خارقة" للحواسيب، لكن المشكلات والإمكانات اللازمة للتحسين هي أيضاً موضع إدراك الباحثين وهذا ما تؤكده المعطيات، عندما يتم اكتشاف أنّ الصور التي تدور فيها الأجسام بطريقةٍ غير معتادة، قد أسيء تصنيف بياناتها، أو أنّ برامج التعرّف على الوجه التي تمّ إنشاؤها أقل قدرة من التعرّف على النساء ذوي البشرة الداكنة بالمقارنة مع ذوي البشرة الفاتحة، وفي هذه الحالات، فإنّ الأمر يتعلّق عادةً بتحسين مجموعات بيانات التدريب المتحيّزة حتى يصبح بالإمكان التعويض عن العجز في تدريب البيانات وتصنيفها، والحقيقة هو أنّ التعلّم الآلي يفشل تمامًا عندما يتعلّق الأمر بالتعرّف على العناصر الموجودة في الصور الغير واضحة المعالم، ورغم أنّ البشر

قادرون على إدراك الخفايا الجمالية، أو الشذوذ، أو السياق التاريخي وحتى السياسي للصور، وذلك بسبب خلفية معارفهم الثقافية، فإنّ برامج التعرّف على الصور تخفق في حالاتٍ مشابهة، لأنه يتعذر الحصول على عددٍ كاف من تصنيفات الصور للتدريب، بالإضافة إلى ذلك، تختلف الطريقة التي يتمّ بها إدراك الأشياء أو الأبعاد المعروضة في الصور وفقًا للخلفية الثقافية أو الاجتماعية للمشاهد، وبالتالي، ستظل هناك قيود قائمة دائمًا، فيما يتعلّق ببرنامج التعرف على الصور وتصنيف بياناتها.

4.3.3 إشكالية الذكاء الاصطناعي مع الاستثناءات

هي قضية أخرى تضاف إلى تحديات الذكاء الاصطناعي، وعلى سبيل المثال يعمل برنامج التعرّف على الصور على الوجه المحدَّد بشكلٍ جيد، ولكن تكمن إشكالية فيما يتعلّق بالصور التي تحوي على عناصر زخرفية، حيث لا يمكن لهذه الصور أن يتمّ تضمينها في مجموعات بيانات التدريب، و في هذا الإطار، يمكن العثور على العديد من الأمثلة، والتي يصف فيها برنامج التعرّف على الصور عناصر خاطئة، ومعظم هذا النوع من الصور تشترك في كونها تصوّر مشاهد غير مألوفة، ولا يمكن فهم معانيها أو جمالياتها أو آثارها، إلّا من خلال خلفية ثقافية معينة، وعليه، فإنّ معرفة السياقات المعقّدة حول الظروف الثقافية، النفسية، وحتى الاقتصادية ليست متوفرة حتى الآن في برامج الذكاء الاصطناعي، ونذكر مثال آخر وهو حالة الأوهام البصرية في الصور46، وهذا أيضاً لا يمكن لأجهزة الحاسوب التعرّف عليه وخاصةً أنها عاجزة عن إنشاء هذه الأوهام في الصور واستخدامها، ورغم أنه لا توجد مشكلة في توليد الوجوه الاصطناعية، أو مقاطع الفيديو، أو الكلام وغيرها،.. إلا أن الحواسيب تفشل في توليد أوهام بصرية اصطناعية، والسبب وراء هذا، هو أن الأجهزة ببساطة لا "ترى"، فالنظر إلى البشر يختلف اختلافًا جوهريًا عن الطريقة التي تعمل بها الرؤية الحاسوبية.

5.3.3 بناء تطبيقات آمنة يكاد يكون مستحيلًا

حتى الآن، ليس من الممكن العثور على جميع الشوائب (Bugs) في رموز البرمجة، سواء كان ذلك بمساعدة البشر أو عبر الوسائل التكنولوجية، وهذا يعني أنّ تقنيات الذكاء الاصطناعي ستكون عرضة لأشكالٍ معينة من الهجمات الحاسوبية بطريقةٍ أو بأخرى، فمن ناحية يتم استخدام التطبيقات الذكية نفسها بشكلٍ متزايد لمنع الهجمات واكتشافها واتخاذ اجراءاتٍ دفاعية تحول دون وقوعها، ومن الأمثلة على ذلك نذكر برنامج "Watson for Cyber Security"الذي يتبع شركة "IBM" أو "For-tiguard Artificial Intelligence"من شركة "Fortinet" ومن ناحيةٍ أخرى، يمكن أن تعمل ذات التقنية على تحليل وكشف طرق الدفاع السيبراني وحتى الالتفاف عليها.

لكن، وبمعزلٍ عن سباق التسلح الذي نشهده في مجال أمن تكنولوجيا المعلومات بين الدول، فإنّ التطبيقات القائمة على الذكاء الاصطناعي لها نقاط ضعف مختلفة في حد ذاتها، ويمكن أن تنطوي نقاط الضعف هذه، على مخاطر جسيمة، ومن الأمثلة على ذلك، نذكر حالات هجمات "تسميم" البيانات حيث يتم تخريب

بيانات التدريب للخوارزميات، وهناك أيضاً قيام المهاجمين بوضع مدخلاتٍ بيانية زائفة قد تؤدي إلى الإساءة، كما أنه من الممكن خداع برامج معالجة النصوص عن طريق إضافة مقتطفات نصية صغيرة إلى النصوص، وبالتالي، يؤدي هذا التلاعب البسيط إلى "فهم" خاطئ تمامًا للنصوص، والشيء نفسه ينطبق على أنظمة التعرّف على الصوت أو الوجه أو الصورة، حيث أنّ التعديلات الطفيفة لملفات الصوت أو الصور تؤدي إلى تفسيرات خاطئة تمامًا، كما أنّ هناك تهديد آخر يتمثّل في سرقة النماذج لإعادة بناء أساليب التعلم الآلي المستخدَمة بالفعل، وذلك من أجل التلاعب بالنظام فيما بعد، باختصار، لا يستطيع الذكاء الاصطناعي جعل العالم أكثر أمانًا فيما يتعلّق بأمن المعلومات ومنع وقوع هجماتٍ إلكترونية.

4 الخاتمة:

سيذكر هذا الجيل زمننا هذا بأن عصر التغيرات الهائلة، فقد تحولنا في ظرف عقود من مجتمع يعتمد بعد الثورة الصناعية على الآلات إلى الاعتماد على المعلومات، وتقدم بنا صريعًا هذا العصر الى لغة حديثة، أكثر أفة وحميمية مجبرًا على التعامل مع نظم خوارزمية مبنية على البيانات، تساعدنا في اتخاذ القرارات، تعتد آلية التعلم بخوارزمية مؤتمتة في شكل روبوتات أو محركات تقديم توصيات. لعل معالجة اللغات الطبيعية أحد المجالات الواسعة متعدد التخصصات التي يربط فيها مهندسي الذكاء الاصطناعي بين علم الكمبيوتر واللسانيات والرياضيات وعلم النفس. ولهذا المجال أهداف بحثية عديدة نظرية وتطبيقية من بينها نمذجة كيفية اكتساب الإنسان للغة إلى برمجة تطبيقات ذكاء اصطناعي كالترجمة الآلية والتعرف الآلي على الكلام والمساعِدات الآلية أو الشات روبوتات مثل "سيري" و"ألكسا"، وجزء كبير من هذا المجال يعمل على ما يسمى بالتقنيات التمكينية كالتحليل الصرفي الآلي مثلا الذي يُستخدم كعنصر في بناء التطبيقات الذكية. أن الأساليب الأكثر استخداما فيه اليوم كما أسلفنا توضيحه في متن هذه الدراسة هي أساليب التعلم الآلي العميق والإحصائي والتي تعتمد بشدة على موارد البيانات. فمثلا، لبناء نظام ترجمة آلي بين لغتين يلزمنا الآلاف بل حتى الملايين من الجمل المترجمة في اللغتين ‘‘أو ما يسمى بالمدونات المتوازية’’ والتي تبني خوارزميات التعلم الآلي منها نماذج حاسوبية للترجمة، وبالطبع، فإن جودة النتائج تعتمد على نوعية الخوارزميات وكمية البيانات المتوفرة لها. وعلى الرغم من كون الخوارزميات والأساليب المستخدمة في معالجة اللغات الطبيعية حيادية لغويا، إلا أننا نجد أن جودة التطبيقات أعلى في لغات معينة كالإنجليزية والألمانية لكونها غنية بموارد البيانات وليس لأي خصائص لغوية تميزها عن لغات فقيرة نسبيا بموارد البيانات كاللغة العربية ولهجاتها.

لغة المجتمع العربي اليوم قد يتخللها تنقل بين العامية والفصحى ومفردات على درجات مختلفة من التعرّب من لغات أخرى كالإنجليزية والفرنسية. أن تطبيقات الذكاء الاصطناعي اللغوية يجب أن تُصمم لخدمة هذا الإنسان واحتياجاته اللغوية كلها: مثلا، أنظمة المساعدة الآلية يجب أن تتعرف تلقائيا على لهجته، وتتفهم أوامره وأسئلته، وتجيب عليها باللهجة التي هو يفضلها كمستخدم. ولدى مطوري تطبيقات الذكاء الاصطناعي قناعة أن

تحقيق ذلك اليوم يعتبر من جانب تقني وتكنولوجي خيالاً علميا" فهناك عديد التحديات التي تستوقفهم، تستدعي من المحتوى الرقمي العربي تعزيز تواجده بداية عبر الانترنت، وتغدية البيانات في شتى مناحي الحياة، وتكوين الكفاءات في مجال الذكاء الاصطناعي، مع مراعاة تمثيل كلا الجنسين ومختلف اللهجات والثقافات. ومراقبة الآلة باستمرار لضمان الحيادية وعدم التحيز، لعل الحديث عن تحليل المحتوى العربي الرقمي يحتاج الى نقاش كبير ينطلق من إرادة بحثية وفترة زمنية للوصول الى توصيات تخدم اهدافه.

5 قائمة المراجع:

المراجع باللغة العربية:

- د.ر غاريسون، تيري اندرسون، و محمد رضوان الأبرش. (2006). *التعلم الإلكتروني في القرن الواحد والعشرون: إطار عمل للبحث والتطبيق* (الإصدار 1). المملكة العربية السعودية: مكتبات ونشر العبيكان.
- طلال أبو غزالة. (2019). *العالم المعرفي المتوقد* (الإصدار 2). المملكة الأردنية الهاشمية: طلال أبو غزالة للترجمة والتوزيع والنشر.
- عبدالهادي السيد عبده. (2021). *المعرفة بين الإنفعال والأخلاق*. مصر: مكتبة الانجلو المصرية.
- محمد محمد الهادي. (2021). *الذكاء الإصطناعي: معالمهوتطبيقاته وتأثيراته التنموية والمجتمعية* (الإصدار 1). مصر: الدار المصرية اللبنانية.

المراجع باللغة الإنجليزية

- David, R. (2020, MARCH 24). The Role of AI in Ensuring Data Privacy. Retrieved from CMSWire: https://www.cmswire.com/information-management/the-role-of-ai-in-ensuring-data-privacy/
- CASILL, A. (2017). Digital labor studies go global: toward a digital decolonial turn. International Journal of Communication, 11, 3934–3954.
- Content Analysis/. (s.d.).
- danah, b., & Kate, C. (2012, May 10). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, 15(5). doi:https://doi.org/10.1080/1369118X.2012.678878
- Duwairi, R., & El-Orfali, M. (2014, May 12). A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. Journal of Information Science, 40(4), 501-513. doi:https://doi.org/10.1177/0165551514534143
- Emanuelle, B., Judy, G., Sven, K., Kuipers, B., Nicholas, M., & Toby, W. (2017). Ethical Considerations in Artificial Intelligence Courses. (P. P. Network, Ed.) AI Magazine, 38(2), 22-34. doi:10.1609/aimag.v38i2.2731
- Gitelman, L. (2013). "Raw Data" is an oxymoron (10 ed.). (M. The MIT Press Cambridge, Ed.) London, England : The MIT Press Cambridge, Massachusetts.

- Hyysalo, S. (2006, Aug). Representations of Use and Practice-Bound Imaginaries in Automating the Safety of the Elderly. Social Studies of Science, 599-626. doi:https://www.jstor.org/stable/25474461
- Khan, B., Gawalt, J., & CooK, F. (2016). Science & engineering indicators. USA: National Science Foundation.
- Kitchin, R. (2017). Thinking critically about and researching algorithms. INFORMATION, COMMUNICATION & SOCIETY, 20(1), 14-29. doi:https://doi.org/10.1080/1369118X.2016.1154087
- Krippendorff, K. (2018). Content Analysis: An Introduction to Its Methodology (4 ed.). USA: SAGE Publications.
- Lundberg, S., & Stearns, J. (2018). Women in Economics. Germany: IZA Institute of Labor Economics.
- Maria, L. (2006). Looking into the future: surveillance, globalization and the totalitarian potential. In L. David, Theorizing Surveillance (1 ed., p. 26). London: Willan. doi:https://doi.org/10.4324/9781843926818
- Markow, W., Hughes, D., & Miller, S. (2017). how the demand for data science skills is disrupting the job market. BOSTON, USA: IBM, BHEF, Burning Glass Technologies.
- McPherson, S. (2018). Artificial Intelligence: Building Smarter Machines (Twenty-First Century Books ed.). (illustrée, Ed.)

مواقع الأنترنت:

- Internet World Stats. (2020, March 31). INTERNET WORLD USERS BY LANGUAGE. Retrieved May 30, 2022, from Internet World Stats - Web Site Directory: https://www.internetworldstats.com/stats7.htm
- Metz, C. (2017, Oct 22). Tech Giants Are Paying Huge Salaries For Scarce AI Talent. Retrieved May 30, 2022, from The New York Times: https://www.nytimes.com/2017/10/22/technology/artificial-intelligence-experts-salaries.html