الدور الآلي للذكاء الاصطناعي في معالجة المستوى الصرفي للغة العربية باستخدام الخوارزميات

The automated role of artificial intelligence in processing the morphological level of the Arabic language using algorithms

عبد الغاني سي موسى[1]، زهرة سعد الله[2]

[1]جامعة وهران 1 أحمد بن بلة (الجزائر)، simoussa.abdelghani@edu.univ-oran1.dz

مخبر اللسانيات وتحليل الخطاب

[2]جامعة وهران 1 أحمد بن بلة (الجزائر)، saadallahchahra@gmail.com

تاريخ الاستلام: 2024/01/20 تاريخ القبول: 2024/10/16 تاريخ النشر: 2024/12/01

ملخص: تهدف هذه الدراسة إلى إبراز الدور المهم الذي يلعبه الذكاء الاصطناعي في معالجة المستوى الصرفي للغة العربية باستخدام خوارزميات التحليل والتوليد، ومن خلال ذلك تم التوصل إلى مجموعة من النتائج أبانت عن الاستفادة الحقيقية التي حققتها لغتنا العربية من هذه التكنولوجيا، إلا أن هذه الاستفادة تبقى جزئية تتحكم فيها عوامل هي بمثابة تحديات يمكن تجاوزها باتباع مجموعة الأسس المقترحة في نهاية الدراسة.

كلمات مفتاحية: الذكاء الاصطناعي؛ المعالجة الآلية للغة؛ الخوارزميات؛ التحليل الصرفي؛ التوليد.

Abstract:

This study aims to highlight the significant role that artificial intelligence plays in processing the morphological level of the Arabic language using analysis and generation algorithms, And through that, a set of results were reached that showed the real benefit that our Arabic language achieved from this technology. However, this benefit remains partial and is controlled by factors that are challenges that can be overcome by following the set of principles proposed at the end of the study.

Keywords: artificial intelligence; natural language processing; algorithms; morphological analysis; morphological generation.

المؤلف المرسل: عبد الغاني سي موسى، **الإيميل**: simoussa.abdelghani@edu.univ-oran1.dz

1. مقدمة:

الذكاء الاصطناعي هو أحد العلوم التي نتجت عن الثورة التكنولوجية المعاصرة والبحث فيه عمل جماعي بالدرجة الأولى يحتم تعاون علماء ومختصين في الحاسب الآلي وعلم اللغة، والمنطق، والرياضيات وعلم النفس(بونيه، 1993، صفحة 13)، كما أن ميدانه البحثي رحب له أبعاد اقتصادية فائقة تظهر آثارها على أصعدة مختلفة وفي ميادين شتى تشمل الصناعة، الاقتصاد ،الطب، الهندسة، والتدريس... ولا يزال الأمل قائما لدى المعنيين بمستقبل الآلة في الوصول بالذكاء الاصطناعي إلى مستويات متقدمة يمكن معها توفير كثير من الجهود والطاقات البشرية من خلال تنمية قدرات الآلة وجعلها قادرة على محاكاة ذكاء الإنسان، ولعل الطفرة الهائلة التي نشهدها اليوم تعجل بتحقيق هذه الآمال والطموحات(عطية، 2019، صفحة 13).

تتفرع عن الذكاء الاصطناعي مجالات معرفية عديدة، منها ما يعرف بمعالجة اللغات الطبيعية؛ الذي يتعامل مع طرق تحليل ونمذجة وفهم اللغة البشرية بصورتيها الرئيسيتين المكتوبة والمنطوقة عبر مستوياتها المتعددة ومعالجة وحداتها في هذه المستويات تحليلا وتوليدا وخلق بيئة تفاعلية قادرة على تحقيق التواصل بين الإنسان والآلة(عطية، 2019، صفحة 14).

واللغةُ العربية إحدى اللغات الطبيعيَّة التي تحظ بعناية الباحثين في مجال الذكاء الاصطناعي عموما ومعالجة اللغات الطبيعية على وجه الخصوص، ذلك أنها واحدة من أكثر اللغات انتشارا في عالمنا المعاصر حيث تأتي في المرتبة الرابعة من حيث عدد مستخدميها،بعد الصينية والأردية والهندية والإنجليزية، ومن شرفة ما سبق تأتي هذه الدراسة الموسومة بـ: الدور الآلي للذكاء الاصطناعي في معالجة المستوى الصرفي للغة العربية باستخدام الخوارزميات لتحاول الوقوف على ما يمكن أن يقدمه الذكاء الاصطناعي للغة العربية في ميدان المعالجة الآلية للغات الطبيعية على المستوى الصرفي من خلال الإجابة على السؤال الرئيسي التالي: هل أثبتت خوارزميات الذكاء الاصطناعي فاعليتها في معالجة وتحليل المستويات اللغوية للغة العربية لاسيما المستوى الصرفي؟ وللإجابة عن السؤال الرئيسي ينبغي علينا الإجابة عن مجموعة الأسئلة الفرعية التالية: ما مفهوم المعالجة الآلية للغات الطبيعية؟ كيف تتم؟ وما هي أهم تطبيقاتها؟ ما المقصود بالخوارزميات؟ وما هي

خصائصها وطرق كتابتها؟ كيف تتم معالجة اللغة العربية آليا على المستوى الصرفي؟ ما هي أهم تحديات المعالجة الآلية للغة العربية على المستوى الصرفي؟ ما هي أهم الحلول التي يمكن من خلالها تجاوز هذه التحديات؟ ولكن قبل التطرق إلى هذه النقاط بالتفصيل، ينبغي الوقوف عند الإطار النظري للذكاء الاصطناعي من خلال تحديد مفهومه، وأهم فروعه الرئيسية.

2.ماهية الذكاء الاصطناعي وفروعه الرئيسية:

ترجع جذور البحوث الخاصة بالذكاء الاصطناعي إلى الاختبار الذي وضع فرضياته العالم الانجليزي ألان تورنج، والذي وصف في الثلاثينات آلة خيالية يمكنها تحديد المشكلات وحلها وكتابة الرموز وقراءتها وتعمل بمقتضاها من تلقاء نفسها(بسيوني، 1994، صفحة 17)،وعلى هذا الأساس فإن كلمة الذكاء الاصطناعي قد تشير إلى الآلة أو الحاسبات على وجه الخصوص، و بالتالي يمكن تعريفه بأنه جزء من علوم الحاسب يهدف إلى تصميم أنظمة ذكية تعطي نفس الخصائص التي نعرفها بالذكاء في السلوك الإنساني(ويتباي، 2008، صفحة 15)؛ ومن خلال التعريف السابق يتضح أن الذكاء الاصطناعي علم يهدف إلى جعل الحاسوب وغيره من الآلات تكتسب الذكاء، وتكون قادرة على أداء مختلف المهام الإنسانية مثل: التفكير، والتعلم، والإبداع، والتخاطب...

للذكاء الاصطناعي فروع عديدة كانت سببا في انقسام العلماء إلى فئتين: الأولى التي تراه كمجال هندسي، والثانية التي تعتبره مجالا علميا بحتا؛ فكل فئة كانت ترى فروع الذكاء الاصطناعي من منظور مختلف، غير أن الجميع كان متفقا على أن تطبيقاته تتعدى المجالات العلمية والهندسية، وفي هذه الدراسة سنكتفي بتغطية الفروع العامة التي تنحصر فيما يلي:

1.2 الأنظمة الخبيرة:

وهي برامج تحتوي على كمية هائلة من المعلومات التي يملكها الخبير الإنساني في حقل معين من حقول المعرفة لكي تقوم بمهام متعلقة بالخبرة البشرية والتي عادة ما تكون من اختصاص البشر، وذلك من خلال مجموعة من الخوارزميات، وبحوث العمليات لاتخاذ الحكم والقرار المناسب(جمعه، 2018، صفحة 198).

2.2 التعلم الآلي:

التعلم الآلي هو فرع من فروع الذكاء الاصطناعي يتعامل مع تصميم وتطوير الخوارزميات التي تسمح لأجهزة الكمبيوتر بتطوير السلوكيات بناء على البيانات التجريبية، مثل بيانات الاستشعار أو قواعد البيانات، حيث ينصب التركيز الرئيسي للدراسة في هذ الفرع على جعل النظام يتعلم تلقائيا من البيانات والبيئة المتاحة وتدريب نفسه من أجل اتخاذ قرارات ذكية، وبذلك فإنه يساعد في إنشاء نظام يصعب بناؤه يدويا بسهولة كبيرة وبجهد أقل(موسى و أحمد، 2019، الصفحات 24-25)

3.2 الروبطة:

وهي من الحقول المتميزة في الذكاء الاصطناعي تهتم بصناعة آلات كهروميكانيكية تتكون من هياكل مشابهة للإنسان، يمكن برمجتها لتؤدي بعض الأعمال الشاقة والمرهقة والخطرة التي يقوم بها الإنسان يدوياً بقوة أكبر وأداء أسرع دون كلل أو تعب وبطريقة آمنة، هذه الروبوتات تعتمد على أنظمة الذكاء الاصطناعي لكي تمتلك القدرة على الحركة، وفهما لمحيط والاستجابة لعدد من العوامل الخارجية(موسى و أحمد، 2019، صفحة 28).

4.2 الرؤية بالكمبيوتر:

ويقصد بها تزويد الكمبيوتر بأجهزة استشعار ضوئية بحيث تمكنه من التعرف على الأشخاص أو الأشكال الموجودة، وذلك عن طريق تطوير عدة أساليب فنية لتحليل الصورة وتمييز الوجوه، فالهدف من مجال الرؤية بالحاسب هو جعل الآلة قادرة على رؤية الوسط المحيط بها والتعرف عليه(مُحَمَّد و أسماء، 2019، صفحة 26).

5.2 معالجة اللغات الطبيعية:

تُعَرَّف معالجة اللغات الطبيعية بأنها: تطويع اللغة بكل تعقيداتها لثنائية الصفر والواحد في برمجيات الحاسب ونظمه بالاعتماد على مجموعة من الوسائل الرياضية التي تمثل حلقة الوصل بين اللغة والحاسوب ومن تلك الوسائل الخوارزميات التي تمثل لغة التخاطب التي يفهمها الحاسوب(الغرابي، 2020، الصفحات 209-214)، فهي إذن مجال فرعي من اللغويات وعلوم الكمبيوتر وهندسة المعلومات

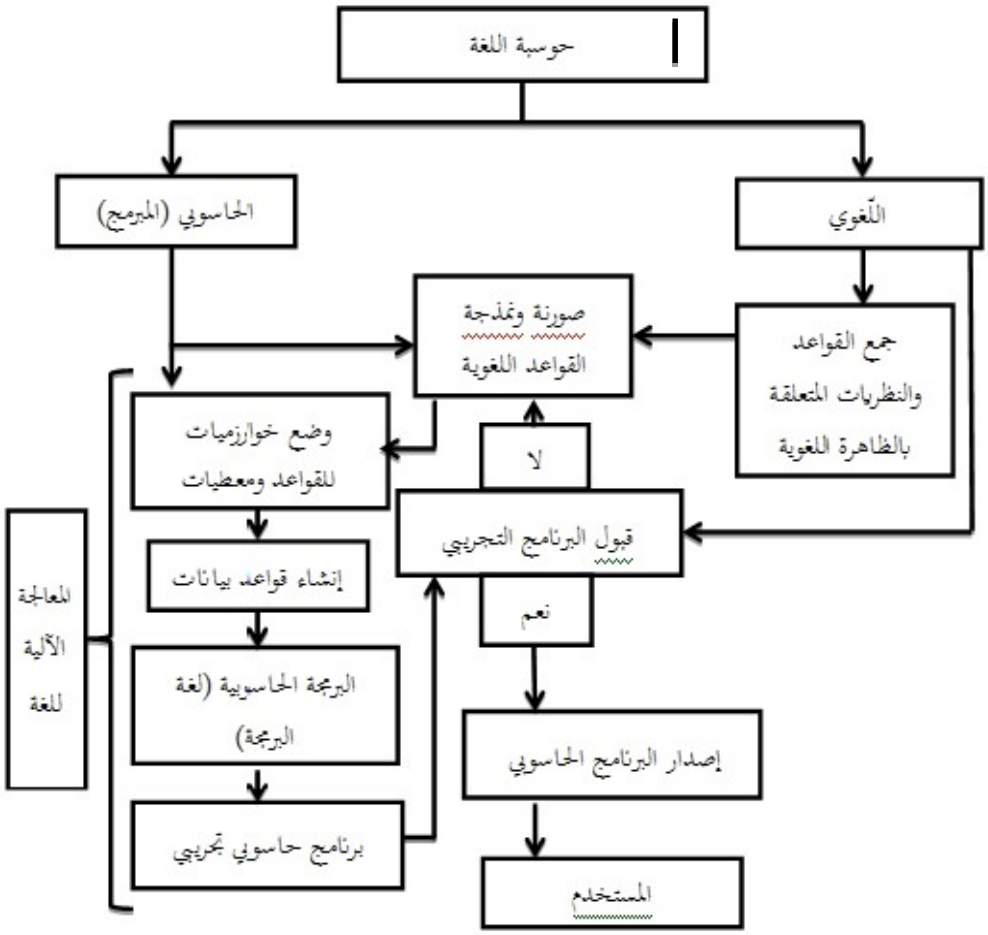
والذكاء الاصطناعي، إذ تهتم بالتفاعلات بين أجهزة الكمبيوتر واللغات البشرية (الطبيعية)، ولا سيما كيفية برمجة أجهزة الكمبيوتر لمعالجة كميات كبيرة من اللغة الطبيعية المنطوقة او المكتوبة وتحليلها إلى بيانات وتوليدها أيضا(Patel & uppili, 2021, p. 4)، وبهذا نخلص إلى أن معالجة اللغات الطبيعية هي جهود تحاول كسر الحاجز بين لغة الإنسان والحاسوب؛ أي بين اللغة التي يستخدمها الإنسان في ظرف عادي طبيعي، ولغة الحاسوب التي يستعين بها لمعالجة وتخزين المعلومات.

إن اللغة العربية قابلة للمعالجة الآلية وبذلك فهي ترتبط ارتباطا وثيقا بالذكاء الاصطناعي فهي مهيأة للوصول إلى مستويات متقدمة للغاية إذا توافرت لها الموارد والطاقات من ناحية، وإذا توفر لها الدعم المناسب من ناحية أخرى، ذلك أن اللغة العربية قياسية في شطر كبير منها، لاسيما في بنيتها الصوتية وقوانينها الصرفية، وفي الوقت ذاته فالعربية لغة معربة وتتمتع بنظام اشتقاقي توليدي، وتتمتع أيضا بنظام كتابي خاص؛ وهي أمور تمثل إشكالات ليست هينة، ولن تكون معالجة هذه الإشكالات يسيرة ما لم تتوافر السبل والأدوات المؤدية إلى ذلك(عطية، 2019، صفحة 14).

3. كيف تتم عملية المعالجة الآلية للغات الطبيعية؟

تتطلب المعالجة الآلية للغة الطبيعية حوارا جادا بين المهندسين وبين اللغويين، فيقدم اللغوي كل تفاصيل الظاهرة اللغوية المراد حوسبتها، وبدوره المهندس (المبرمج) يترجم تلك التفاصيل إلى خوارزميات (صالح، 2012، صفحة 84)،ومن ثم يقوم بإدخالها إلى الحاسوب، وكما هو معلوم أن الحاسوب في عملياته الداخلية يتعامل مع الأرقام الثنائية (0.1) وهو لا يتعامل مع أي شيء آخر، لذا يجب تحويل تلك الخوارزميات أيضا إلى لغة يفهمها الحاسوب عن طريق استدعاء مجموعة من الكلمات المفتاحية من إحدى لغات البرمجة مرتبطة مع مجموعة من القواعد المحددة حول كيفية تجميع هذه الكلمات لتصبح عبارات يمكن للمفسر أو المترجم ترجمتها إلى لغة الآلة الثنائية(Swinnen, 2012, p. 5)، ومن خلال هذه الوسائط المتكاملة تتم معالجة اللغات الطبيعية، ومن ثم تصير الآلة قادرة على التفاعل مع الإنسان بالتواصل معه عبر وسائل التعبير المختلفة كالكلام، الصوت، الصورة، الفيديو، النصوص...(جمعه،

2018، صفحة 199)، والمخطط التشجيري التالي يوضح ويبسِّط ما قيل حول عملية المعالجة الآلية للغة:



المخطط 1: مراحل المعالجة الآلية للغة

1.3 أهمية المعالجة الآلية للغة العربية:

تكمن أهمية المعالجة الآلية للغة العربية بشكل خاص في تنظيم الثروة اللغوية في سائر مستوياتها الصوتية، الصرفية، النحوية والدلالية، من خلال بناء قاعدة معلومات ضخمة تساعد على تخزين هذه الثروة، كما تهدف إلى بناء موسوعة إلكترونية للغة العربية تساعد على إخضاعها لمنهجية العلم المضبوط ومطالب المعالجة الآلية الدقيقة، كما تساعد على تصميم برامج صحيحة للترجمة الآلية تراعي خصوصيات اللغة

العربية(منعم، 2015، صفحة 181)، وبذلك يتمكن الإداري والعالم والمسؤول الحكومي من الوصول بسهولة وسرعة إلى المعلومات المطلوبة في أي مكان وزمان،ويتحقق ذلك من خلال التلخيص الآلي وتحليل النصوص بالإضافة إلى البحث الذكي عبر شبكة الإنترنت في الأجيال الجديدة المعروفة بالويب الدلالي؛ ولتوضيح هذا المفهوم تخيّل نفسك في رحلة صيفية إلى الساحل الشمالي وأنت ترغب في متابعة أداء أسهم شركتك في البورصة. كيف ستتعامل مع ذلك؟ ببساطة، ستتحدث إلى نظام السيارة قائلًا: أريد تقريرًا مفصلًا حول أداء أسهم شركة كذا في البورصة خلال اليومين الماضيين. وماذا يحدث بعد ذلك؟ يقوم نظام السيارة بتحويل كلامك إلى أوامر بحث نصية ويرسلها عبر تقنيات الاتصال اللاسلكية المتطورة إلى شبكة الإنترنت التي تفحص بدورها أداء أسهم شركتك خلال فترة محددة وتعيد إرسال المعلومات النصية إلى نظام السيارة. هنا يتم تحويل المعلومات النصية إلى كلام يمكنك سماعه بأذنيك، مما يسمح لك بالتركيز بشكل طبيعي على القيادة(التميمي، 2010).

إن تطبيقات المعالجة الآلية للغات الطبيعية كثيرة ومتنوعة تجعل الآلة قادرة على القيام بالعديد من المهام اللغوية كالترجمة اللغوية، التعرف على الكلام، تلخيص النصوص، استخلاص المعلومات، توليد الكلام، صناعة المعاجم الإلكترونية، استرجاع المعلومات، الإحصاء اللغوي، الفهرسة الآلية، التدقيق الإملائي والنحوي... وغيرها من التطبيقات(حسين، 2021، صفحة 507)، وكما هو معروف أن المعالجة الآلية موجهة لجميع اللغات، وتشمل جميع مستويات اللغة ولا تتحقق إلا بالتجريد والترميز الرياضي؛ والمقصود بذلك توصيف اللغة على شكل خوارزميات رياضية تمثل مادة للاستعمال في الحواسيب وهذا ما تم الإشارة إليه سابقا، لذلك فما المقصود بالخوارزمية وما هي خصائصها وطرق كتابتها؟

4.مفهوم الخوارزمية:

استمد مصطلح الخوارزمية من اسم العالم مُحَمَّد بن موسى الخوارزمي مؤلف كتاب الجبر والمقابلة، وكان هذا الكتاب قد ترجم إلى اللاتينية بعنوان: (Algoritmi de NumeroIndorum) واشتقت المصطلحات المقابلة لعبارة الخوارزمية في جميع اللغات الأوروبية من اسم هذا الكتاب باللاتينية، فكانت (Algorithm) بالإنجليزية، و (Algorithmus) بالألمانية، و (Algorithme) بالفرنسية، ولكن هذا

المفهوم ارتبط في البداية بقواعد إجراء الحسابات باستخدام الأرقام العربية وبعد ذلك نقل المصطلح من مجاله الرياضي ليستخدم في ميدان المعلوماتية، وفي مقاربة أولية لتحديد مفهوم الخوارزمية يمكن تعريفها بأنها مجموعة من العمليات القابلة للتنفيذ محدودة العدد ومرتبة على شكل خطوات متسلسلة تهدف لحل مشكل ما، وهي بمثابة لغة الحوار مع الحاسوب تسهل عليه عملية تحليل المدخلات لكونها أسرع في عملية التنفيذ(المنجد، الآلة الذكية من ديكارت وحتى دماغ غوغل، 2020، صفحة 117)وبناء الخوارزمية هو المرحلة الأولية لتصميم أي نظام حاسوبي؛ إذ يعتمد عليها في بناء أنظمة معالجة اللغات الطبيعية باستعمال لغة برمجة معينة(الغرابي، 2020، صفحة 210).

1.4 خصائص الخوارزمية:للخوارزمية خمس ميزات مهمة:

المحدودية: يجب أن تكون الخوارزمية محدودة بقدر ما يجب أن تنتهي بعد عدد محدود من الخطوات، وهذا مهم بشكل خاص في الخوارزميات التكرارية التي تقوم بالتقريب المتتالي في محاولة حل مشكلة بمستوى عال من الدقة(Jurs, 1996, p. 9)، وكمثال على ذلك عملية قسمة العدد 10 على 3 التي تكون غير محدودة ولا يسمح بها داخل البرنامج.

الوضوح:يجب تحديد كل خطوة في الخوارزمية بدقة، والإجراءات التي سيتم تنفيذها بشكل لا لبس فيه (Juul, 2011, p. 62)، ومثال ذلك عبارة « Add 6 or 7 to X » التي تعتبر غير واضحة ولا يسمح بها داخل الخوارزمية.

المدخلات: تحتوي الخوارزمية على صفر أو أكثر من المدخلات التي يتم تقديمها في البداية قبل بدء الخوارزمية، أو ديناميكيا أثناء تشغيل الخوارزمية، ويتم أخذ هذه المدخلات من البيئة التي تعمل فيها.

المخرجات:تحتوي الخوارزمية على واحد أو أكثر من كميات المخرجات التي لها علاقة محددة بالمدخلات.(Papazian, 2022, p. 196)

الفعالية:من المتوقع أن تكون الخوارزمية فعالة بشكل عام، وهذا يعني أن تكون كل خطوة فيها ممكنة الحل وينبغي أن تجري في فترة زمنية محدودة(Juul، 2011، صفحة 62)، فمثلا عبارة 0/3 لا يمكن حلها أبدا.

2.4 طرق صياغة الخوارزمية:هناك عدة طرق لكتابة الخوارزميات، نجملها فيما يلي:

استخدام اللغات الطبيعية:

يمكن كتابة الخوارزمية باللغة الطبيعية البشرية (العربية، الفرنسية، الإنجليزية...)، وميزة اللغة الطبيعية هي أن كل شخص يفهم اللغة يمكنه قراءة الخوارزمية وكتابتها، ومع ذلك غالبا ما تفتقر اللغات الطبيعية إلى الدقة التي تحتاجها عملية كتابة الخوارزميات، كما أن أجهزة الكمبيوتر تواجه مشاكل في فهم اللغات الطبيعية.(Itmazi, 2017, p. 41)

استخدام لغات البرمجة:

لكي يتمكن الكمبيوتر من تنفيذ الخوارزمية المقدمة له بشكل جيد يتطلب الأمر التعبير عنها بإحدى لغات البرمجة كال C++ أو Java...، فهذه اللغات توفر الدقة اللازمة التي تكون بسيطة بما يكفي لفهمها وتفسيرها لأجهزة الكمبيوتر.(Afriyie, 2012, p. 342)

استخدام الشفرة الزائفة:

الشفرة الزائفة هي مزيج من اللغة الطبيعية ولغة البرمجة، والخوارزمية المكتوبة بلغة زائفة أكثر دقة من خوارزمية مكتوبة بلغة طبيعية، ولكنها أقل دقة من خوارزمية بلغة برمجة، ومن ناحية أخرى غالبا ما تكون خوارزمية الشفرة الزائفة أسهل في القراءة والكتابة من تلك التي يتم التعبير عنها بلغة البرمجة، كما أنها تقوم بتوجيه المبرمجين لتجنب الأخطاء الدلالية أثناء عملية الترميز(Afriyie, 2012, p. 342)

استخدام المخططات الانسيابية:

لزيادة وضوح الخوارزمية يُعبر عنها عادة بمخطط يتم من خلاله ملاحظة الخطوات وتحديد الأخطاء فيها وبالتالي تصحيحها بسهولة، والمخططات الانسيابية تبنى بواسطة سلسلة من الأشكال الهندسية يحتوي كل منها على شرح وصفي، وتتصل هذه الأشكال بواسطة أسهم، إضافة إلى عبارات باللغة الطبيعية، ويفضل استخدام هذه الطريقة حينما تكون الخوارزمية بسيطة وقصيرة(الجواد و مالك، 2008، صفحة 12).

5. معالجة المستوى الصرفي للغة العربية باستخدام الخوارزميات:

يعد الصرف العربي رابطة العقد لعناصر المنظومة اللغوية فهو ركيزة الفونولوجي ومدخل النحو وأساس تنظيم المعجم، وفوق هذا كله هو الخط الساخن لالتقاء مباني اللغة ومعانيها وهذا ما يؤكد على أهمية معالجته آليا(علي، 1988، صفحة 247)باستخدام خوارزميات التحليل والتوليد لضمان فعالية ودقة عملية الإنتاج اللغوي، ويقصد بالتحليل الصرفي الآلي للكلمة في اللغة العربية ربط كلمات النص بالعناصر الصرفية الأولية التي تدخل في تكوينها، وكذلك بالقيم النحوية دون اعتبار موقعها، فيتم في التحليل الانتقال من الكلمة إلى جذرها الأصلي، أي أن الحاسوب يعالج الكلمات العربية المشكولة جزئيا أو كليا، أو غير المشكولة فيصف ما يطرأ عليها من تغيير سواءً أكانت زيادة أو نقصانا أو إعلالا أو إبدالا أو إدغاما أو قلبا، كما يحدد نوعها وميزانها الصرفي وسابقتها أو سوابقها، ولاحقتها أو لواحقها وحالتها الإعرابية ودلالتها، فإذا احتوت الكلمة المراد تحليلها على حروف غير مشكولة وضع الحاسوب الحركات الممكنة لها اعتمادا على بيانات مخزنة(المجيد، 2020، صفحة 30)، ومن تمثيل ذلك تحليل كلمة "وبإيصاله" صرفيا:

الجدول 1: مثال عن عملية التحليل الصرفي للكلمة العربية آليا

السوابق	الواو: حرف عطف، الباء:	**الميزان الصرفي**	إيعال
جذع الكلمة	إيصال	**الحالة التصريفية**	مفرد، مذكر، مجرور
جذر الكلمة	وصل	**العلامة**	الكسرة الظاهرة
الرتبة النحوية	مصدر ثلاثي مطرد	**اللواحق**	الهاء: ضمير الملكية للمفرد الغائب
الصيغة	إفعال		

المصدر:من إعداد الباحث بالاعتماد على(الشمس، 2020، صفحة 18)

أما التوليد الصرفي الآلي فهو عكس عملية التحليل التي تستخدم لفهم اللغة، إذ يتمثل في ابتكار صيغ جديدة وكلمات مشتقة من المدخلات اللغوية في الحاسوب انطلاقا من الجذر بوصفه نقطة البداية مرورا بالقناة الخوارزمية(مهديوي، اللغة العربية وطالب مجتمع المعرفة الفرص والتحديات والحلول، 2020، صفحة 12)التي تقوم بدمج عناصر الكلمة كالسوابق واللواحق... مع عناصر الجذر الأصلية وتُجري عليها عمليات الضبط الفونولوجي اللازمة من تضعيف وإدغام وحذف...، وبذلك تنتهي الخوارزمية إلى تركيب الكلمة في صورتها النهائية تامة التشكيل وتقدمها للمستخدم في شكل مخرجات(الميهوبي، 2018، صفحة 146).

1.5 خوارزميات التحليل الصرفي:

وهي تمثل الأساس الذي تبنى عليه تطبيقات معالجة اللغة العربية الطبيعية، وتطبيقات تعليم اللغة العربية لغير الناطقين بها، ومحركات البحث ونظم الاسترجاع، ونظم الترجمة الآلية ثنائية الاتجاه... وتعتمد خوارزميات التحليل الصرفي في عملها على فحص الخواص الصرفية والنحوية لقاعدة ضخمة من المفردات العربية المستقاة من تحليل ذخيرة لغوية ضخمة، كما أنها تتميز بتغطية أكبر، وقدرة سريعة على التوسع، وإمكانية تناقلها بين مختلف نظم التشغيل(الفلاح، 2009، الصفحات 615-616)، وتنقسم خوارزميات التحليل الصرفي إلى أربعة أقسام:

خوارزميات التحليل الصرفي القاموسي:وتقوم بتفسير الجمل والكلمات وتبحث عن معانيها انطلاقا من القاموس الملحق بالنظام(عفيفي، 2015، صفحة 29).

خوارزميات التحليل الصرفي اللغوي:وتقوم باستخلاص العناصر الأولية لبنية الكلمة في اللغة العربية وتحدد سماتها الصرفية، والصرف صوتية، والصرف نحوية(الغامدي، 2017، صفحة 51)وذلك من خلال التعامل مع: السوابق الممكنة (Prefix)، الأوزان الأساسية(Patterns)، جميع اللواحق الممكنة (Suffix)، جميع أصناف الجذور، جميع الكلمات الجامدة، الكلمات الشاذة، الكلمات الأعجمية، تحديد قواعد التعدية واللزوم بالنسبة للأفعال، تحديد قواعد التصريف في الماضي والمضارع والأمر في أحوالها المختلفة، تحديد قواعد الإعلال والإدغام والإبدال والهمزة(حماده، 2010، صفحة 365).

ولتتعامل الخوارزمية مع هذه الخصائص تعتمد على مجموعة من اللوائح المتفاعلة فيما بينها والتي تم إدخالها سابقا في قاعدة بيانات الجهاز، هذه اللوائح تحتوي على جميع السوابق واللواحق والجذور والأوزان العربية وغيرها من اللوائح الأخرى، فكلمة "فيصيبها" تحلل على النحو التالي:

الجدول 1: مثال عن عملية التحليل الصرفي للكلمة العربية آليا

الكلمة المشكولة	السابق	نوع الكلمة	وزن الكلمة	جذر الكلمة	الحالة الإعرابية	جذع الكلمة	اللاحق
فَيُصِيبُهَا	الفاء	فعل مضارع مبني للمعلوم	يفيل	صوب	ثلاثي مزيد منصوب مسند إلى الغائب (هو) متعد	يصيب	ها

المصدر: من إعداد الباحث بالاعتماد على(مولوج، 2020، صفحة 72)

خوارزميات التحليل الصرفي التبادلي:ويُعتمد عليها للوصول إلى ما يعرف بالمادة اللغوية؛ أي الجذر اللغوي للكلمة المدخلة في الجهاز، حيث تقوم هذه الخوارزمية التي تم تدريبها مسبقا بإجراء عمليات الاشتقاق لكلمة معينة ومن ثم ملاحظة واستخراج الحروف الأصلية للجذر التي تكررت في جميع عمليات الاشتقاق ولم تحذف إلا في حالة الإعلال أو الإبدال، ومثال ذلك كلمة (الاجتماع) جذرها (ج م ع) بدليل تكرار هذه الأحرف الثلاثة في جميع الكلمات المشتقة منها، مثل: جَمَعَ، يَجْمَعُ، جَامِعٌ، مَجْمُوعٌ، مُجْتَمَعٌ، جَامِعَةٌ، تَجَمُّعٌ، ...(سليمان، 2016، صفحة 12).

خوارزميات التحليل الصرفي المصدري:ومن مهام هذا النوع من الخوارزميات التصحيح الآلي للأخطاء الإملائية والنحوية الواردة في الكلمات والنصوص(رمضان، 2019، صفحة 192)، وفحص انتماء جذور الكلمات المدخلة إلى معجم الخوارزمية فما وجدته ينتمي إلى معجمها أبقته وما لم تجده استبعدته كما تعمل على معالجة شكل الكلمة ووضع الحركات على حروفها إضافة إلى فهم وحفظ واسترجاع الكلمات العربية وكذا استخلاص العلاقة بين كلمات الجمل والعبارات(الألوسي، 2015، صفحة 94) وللقيام بهذه المهام ينبغي أن يتوفر الحاسوب على ذخيرة لغوية ضخمة.

2.5 خوارزميات التوليد الصرفي:

وتهدف إلى تنمية الرصيد اللغوي وإثرائه من خلال إحداث وحدات معجمية جديدة لها صيغتها الصرفية المستقلة ودلالاتها الخاصة اعتمادا على عمليات الاشتقاق والتركيب والنحت والاختزال، فتبدأ من الجذور بوصفها علامة صوتية مجردة ومادة تنظيمية لمداخل المعجم وما يضاف إلى الجذور من سوابق في أولها ودواخل في وسطها وكذلك اللواحق في آخرها وتنتهي بالمخرج؛ أي المفردة المعجمية(خيرة و هدنة ، 2021، الصفحات 1544-1545)، وينقسم التوليد الصرفي إلى نمطين: توليد صرفي لغوي يقوم على توليد المداخل المعجمية من جذورها وفق قواعد مضبوطة، وتوليد صرفي غير لغوي يقوم على توليد الجذور من جذور أخرى، كتوليد الرباعي من الثلاثي أو الثلاثي من الرباعي، وقد أشار القدماء العرب إلى بعض الأدوات الإجرائية التي يتم بواسطتها هذا النوع من التوليد من قبيل القلب، الإبدال، التوسع...(مهديوي،

التوليد الصوري للأسماء البسيطة في اللغة العربية، 2011، صفحة 42)، وتمر عملية التوليد الصرفي بعد خضوع جذر المفردة لسلسلة من التحويلات الصرفية والصوتية عبر المراحل التالية:

إدخال الجذر إلى قاعدة بيانات الجذور، ثم اختيار نوع الجذر (ثلاثي، رباعي، صحيح، معتل)، بعدها ينتقل الجذر من المستوى الصوتي المجرد إلى المستوى المحسوس بعد تطبيق خوارزميات التوليد الصرفي على البنية الصورية للجذر، وذلك بإضافة مورفيمات الزيادة والحركات، ومن ثم تفرغ الكلمة في قالب صرفي (ميزان صرفي) مطابق لها يحدد هويتها بدقة، وبعد الانتهاء من العمليات السابقة يكون الناتج قوائم من المفردات تامة البناء ومرتبة في جداول وخانات على شكل مصفوفات(فيران، 2021، صفحة 471).

6. تحديات المعالجة الآلية للغة العربية على المستوى الصرفي:

معالجة اللغة العربية صرفيا وجعلها ميسرة أمام الآلة يمر عبر تحديات كثيرة نفصل بعضها فيما يلي:

– غياب الحركة الصرفية في بعض الصيغ مما يؤدي إلى غموضها، ومن ثمعجز أنظمة الترجمة على نقلها نقلا صحيحًا كما يحد مثلا بين الفعل ومصدره فيمثل: لَعِبَ ولَعِبٍ، رَكَضَ و رَكضٍ، فَتَحَ و فَتحٍ وغيرها.

– تداخل بعض الأبنية في بعض الأبواب الصرفية كما الحال مع أسماء الفاعل والمفعول في باب مختار كـ (مُكتال، ومُقتاد، ومُبتاع)، وأشكالها لما حدث فيها من إعلال، لأن اسم المفعول يكون فيها على زنة (مُفتَعَلٌ)، واسم الفاعل على زنة (مُفتَعِلٌ)، فالتفريق بينها في مثل هذه الحالات يكون صعبًا على الآلة، مما يحدث لها التباسا في إنتاج المخرجات.

– تنوع أصل اشتقاق بعض الصيغ الصرفية، من ذلك (جائزٌ) التي هي اسم فاعل لـ (جارَ)، وكذلك لـ (جَأَرَ)، و(سائلٌ) التي هي اسم فاعل من (سَأَلَ) و(سَالَ) وما شاكلهما.

– ما يأتي من جموع التكسير دالا على المفرد والجمع جميعا مثل: (فلك) و(هجان) وغيرها، ومن أمثلة ذلك في المفرد مجيء (فلك) دالة على المفرد في قوله تعالى: ﴿في الْفُلْك المشْحون﴾ [سورة الشعراء، الآية 119]ومجيئها دالة على الجمع في قوله تعالى﴿والفلك التي تجري في البحر﴾[سورة البقرة، الآية 164].

– تعدد الوظائف النحوية لبعض الوحدات الصرفية مثل (حاشا) التي تكون اسما،أو اسم فعل، أو فعلا أو حرف خفض في اللغة العربية.

– وغيرها من صيغ المبنى للمعلوم والمجهول، والتذكير والتأنيث، والّتعدي واللزوم...التي لا يمكن للآلة أن تفهمها بسهولة في غياب المعطيات اللازمة فتلتبس عليها،وفي هذه الحالة يسبب اللبس الصرفي لبسًا تركيبيًا ودلاليًا(طيبي، 2018، الصفحات 199-200).

7. الأسس المقترحة لتجاوز تحديات المعالجة الآلية للصرف العربي:

لقد قدم الدكتور نبيل علي في كتابه "اللغة العربية والحاسوب" مجموعة من الأسس العملية للتغلب على مشاكل معالجة الصرف العربي آليا(علي، اللغة العربية والحاسوب دراسة بحثية، 1988، الصفحات 299-301) منها ما يلي:

– ضرورة تعامل المعالج الصرفي الآلي مع أطوار التشكيل المختلفة للنصوص العربية، كما ينبغي أن يكون نظام الصرف الآلي تجزيئيا مكونا من عدة آليات متخصصة تربط بينها علاقات ترابط واضحة ومحددة، ويجب أن تقوم تجزيئية النظام الآلي على أساس لغوي لا فني، والمقصود بذلك تقسيم النظام إلى عناصر تعكس الوظائف الأساسية للمنظومة الصرفية، لا الخطوات الإجرائية للبرنامج.

– يجب أن يتعامل المعالج الصرفي مع ثنائية الصيغة الصرفية والميزان الصرفي (البنية العميقة والبنية السطحية) حيث يكمن في العلاقة الثنائية بينهما قدر كبير من السر الصرفي مع مراعاة اعتبارات التكامل بين المعالج الصرفي والمعالج النحوي.

– يجب أن يكون المعالج الصرفي عاما غير مقيد بشريحة معينة من النصوص، أو بنطاق معين من الموضوعات، ويجب أن يكون شاملا لجميع الجذور المعجمية الممكنة والصيغ الصرفية.

– فصل القواعد الصرفية عن البرنامج الآلي، وذلك لإمكان تعديل هذه القواعد بأقل كمية من التغييرات في برنامج المعالجة.

– مراعاة ثنائية التحليل والتوليد وقدرة المعالج على العمل في الاتجاهين؛ ويقصد بذلك إمكانية تحليل الكلمات إلى عناصرها الأولية أو تركيب هذه العناصر في هيئة كلمات نهائية، وفي هذا الصدد ينبغي الإشارة إلى أن قواعد الصرف العربي في كثير من جوانبها تنحاز نحو التوليد، وعلى مصمم المعالج الصرفي

الآلي أن يسعى لإحداث التوازن الوظيفي وذلك بجعل قواعده وبرامجه أقل حساسية لغرض المعالجة واتجاهها.

– ضرورة تعريض المعالج الصرفي لاختبارات قاسية للتحقق من سلامة الأسس اللغوية المبني عليها، والتأكد من اكتمال تغطيته لجميع الحالات الصرفية والمعجمية للكلمة العربية.

– الاهتمام بشق المعنى الذي يمثل الغاية القصوى للتنظير اللغوي للصرف ومعالجته الآلية.

8. خاتمة:

خوارزميات التحليل والتوليد الصرفي تمثل جزءا أساسيا من تكنولوجيا معالجة اللغة العربية، فهي تتيح للآلة فهما أعمق لقواعد الصرف العربي وتسهل على المبرمجين عملية تصميم تطبيقات ذكية تفاعلية وأنظمة للترجمة الآلية وبرامج لتصحيح النصوص العربية، ومن خلالها أيضا يمكن تعزيز وتسهيل التفاعل بين الإنسان والآلة، لكن تطويرها في الوطن العربي يشهد تأخرا ملحوظا مقارنة بالدول الغربية، وهذا يمكن أن يرجع جزئيا إلى تحديات فريدة تواجه اللغة العربية في هذا المجال، منها ما يتعلق ببيئة البحث اللغوي، ومنها ما يتعلق بخصوصيات اللغة العربية كتنوع اللهجات مثلا، علاوة على ذلك قد تكون البيانات المتاحة لتطوير هذه الخوارزميات غير كافية أو غير دقيقة مما يؤثر على جودة النتائج ويزيد من احتمالية حدوث أخطاء في المخرجات، ولتجنب هذه المشاكل نوصي بالمقترحات التالية:

– تعزيز التعاون البحثي بين المختصين في المجال الهندسي الرياضي الحاسوبي والمختصين اللغويين العرب.

– إجراء حملات توعوية تؤكد على أهمية البحث العلمي في معالجة اللغة العربية آليا عبر مختلف مستوياتها على نطاق المجامع اللغوية العربية وأقسام المعلوماتية في الجامعات العربية، وتوجيه الأبحاث نحو التطبيق العملي.

– تكوين إطارات بحثية في أقسام اللغة العربية ذات معرفة حاسوبية جيدة وفي أقسام المعلوماتية ذات معرفة جيدة باللغة العربية، لكي يتم التواصل والبحث العلمي على أتم وجه.

– العمل على مشروع إعداد معجم محوسب شامل يراعي السياقات اللغوية والمقامية للكلمات داخل النصوص.

توفير الإمكانات المادية اللازمة لمسايرة الركب التكنولوجي في مجال تطبيقات المعالجة الآلية للغة العربية عامة.

– العمل على ترجمة كل ما ينتجه الغرب في مجال اللسانيات الحاسوبية وكل التطبيقات اللغوية من أجل الاطلاع على كل ما هو جديد.

9. قائمة المراجع:

1.9 الكتب العربية:

– آلان، بونيه، (1993)، الذكاء الاصطناعي واقعه ومستقبله، عالم المعرفة، الكويت.

– مُحَمَّد، عطية وآخرون، (2019)، العربية والذكاء الاصطناعي، دار وجوه للنشر والتوزيع، السعودية.

– عبد الحميد، بسيوني، (1994)، الذكاء الاصطناعي ومقدمة برولوج، دار النشر للجامعات المصرية، مصر.

– بلاي، ويتباي، (2008)، الذكاء الاصطناعي، دار الفاروق للاستثمارات الثقافية، مصر.

– مصطفى، حمدي محمود جمعه، (2018)، النظام القانوني للوكالة بالعمولة الإلكترونية، دار الفتح للطباعة والنشر، مصر.

– عبد الله، موسى وأحمد، حبيب بلال، (2019)، الذكاء الاصطناعي ثورة في تقنيات العصر، المجموعة العربية للتدريب والنشر، مصر.

– كريمة، محمود مُحَمَّد وأسماء، السيد محمود، (2019)، تطبيقات الذكاء الاصطناعي ومستقبل تكنولوجيا التعليم، المجموعة العربية للتدريب والنشر، مصر.

– عبد الرحمان، الحاج صالح، (2012)، بحوث ودراسات في اللسانيات العربية، موفم للنشر، الجزائر.

– سناء، منعم، (2015)، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، عالم الكتب الحديث، الأردن.

– مُحَمَّد، بشير المنجد، (2020)، الآلة الذكية من ديكارت وحتى دماغ غوغل، E-kutub Ltd، إنجلترا.

- دلال صادق، مصطفى الجواد و مالك، صالح علي، (2008)، أساليب البرمجة، دار اليازوري العلمية، الأردن.

- نبيل، علي، (1988)، اللغة العربية والحاسوب دراسة بحثية، تعريب للنشر والتوزيع، الكويت.

- خالد، حوير الشمس، (2020)، اللسانيات الحاسوبية تنظيرا وتطبيقا، مركز الكتاب الأكاديمي، الأردن.

- جهاد، عفيفي، (2015)، الذكاء الاصطناعي والأنظمة الخبيرة، دار أمجد، الأردن.

- منصور، بن مُحَمَّد الغامدي، (2017)، مدخل إلى اللسانيات الحاسوبية، دار وجوه للنشر والتوزيع، السعودية.

- صباح، علي سليمان، (2016)، أصول الألفاظ غير الداخلة في القرآن الكريم دراسة تأصيلية دلالية، دار غيداء للنشر والتوزيع، الاردن.

- هاني، إسماعيل رمضان، (2019)، معايير عناصر اللغة العربية للناطقين بغيرها، Akdem للنشر والتوزيع، تركيا.

- تيسير، الألوسي، (2015)، تكنولوجيا تعلم اللغة العربية، مركز الكتاب الأكاديمي، الأردن.

- سلوى، حماده، (2010)، مقترح لمعايير وضوابط تقييم المحللات الصرفية، ورقة بحثية ضمن وقائع الندوة الدولية لعلوم وهندسة الحاسوب، فيليبس للنشر والتوزيع، الولايات المتحدة الأمريكية.

2.9 الكتب الأجنبية:

-Ankur A, Patel and Uppili, Arasanipalai, (2021), natural language processing in the entreprise teaching machines to read write and understand, O'reilly, USA.

- Gérard, Swinnen, (2012), apprendre à programmer avec python 3, eyrolles, France.

- Peter C, Jurs, (1996), Computer software applications in chemistry, John wiley and sond, USA.

- Jesper, Juul, (2011), Half-real video games between real rules and fictional worlds, MIT press, USA.

- Armen V, Papazian, (2022), the space value of money rethinking finance beyond risk and time, palgravemacmillan, England.

- Jamil Ahmed, Itmazi, (2017), fundamentals of computer and programming an Arabic textbook, Phillips publishing, USA.

- Bright, SiawAfriyie, (2012), Concise ICT fundamentals, Trafford publishing, USA.

3.9 المقالات:

– سلطانة بنت مُحَمَّد الغرابي، (2020)، بناء خوارزمية حاسوبية لتوليد جموع القلة في القرآن الكريم، مجلة الدراسات اللغوية، م 22(3).

– عيجولي حسين، (2021)، المعالجة الآلية للغة، مجلة دراسات وأبحاث في العلوم الإنسانية والاجتماعية، م 13(2).

– قدوري عبد المجيد، (2022)، مستويات الحوسبة اللسانية في النص القرآني، مجلة حوسبة اللغة العربية، م 1(1).

– عمر مهديوي، (2020)، اللغة العربية وطالب مجتمع المعرفة الفرص والتحديات والحلول، مجلة حوسبة اللغة العربية، م 1(1).

– عبد الله بن عبد العزيز الميهوبي، (2018)، بناء برنامج حاسوبي لتوليد أفعال اللغة العربية وتصريفها، مجلة مجمع اللغة العربية بمكة المكرمة، ع 17(1).

– عبد الحميد الفلاح، (2009)، سبل توطين التقانة باللغة العربية صناعة تقانة المعلومات نموذجا، مجلة اللغة العربية، م 11(1).

- عمر مهديوي، (2011)، التوليد الصوري للأسماء البسيطة في اللغة العربية، مجلة اللسانيات، م 17(2).

- نجوى فيران، (2021)، خوارزميات الذكاء الاصطناعي ودورها في التحليل الآلي للغة العربية على المستوى الصرفي، مجلة دراسات معاصرة، م 5(2).

- أحمد طيبي، (2018)، حاجة الترجمة الآلية إلى استحضار البعد الدلالي والتداولي، مجلة التعريب، ع 49.

4.9 المداخلات:

- مُحَمَّد أمين، مولوج، (2019)، المعالجة الآلية للغة العربية بين الواقع والمأمول دراسة تطبيقية لحديث إنما الأعمال بالنيات برنامجا مشكال والخليل أنموذجين، اللغة العربية وبرامج الذكاء الاصطناعي الواقع والرهانات، جامعة مصطفى اسطمبولي، الجزائر.

- عيشون خيرة وهدنة خالد، (2021)، التوليد اللغوي في عمقه التراثي وبعده المعاصر، مجلة علوم اللغة العربية وآدابها، م 13(1).

المواقع الالكترونية:

- شاكر التميمي، (2010)، المعالجة الآلية للغة العربية جهود وتحديات، https://www.startimes.com/?t=21168801، (اطلع عليه بتاريخ: 2022/09/03، على الساعة: 15:40)