مجلة العدوي للسانيات العرفنية وتعليمية اللغات
Issn: 2773/ 3300
http://www.asjp.cirect.dz/en/presentation Revue/804
المجلد:05العدد: 01 أكتوبر (2025) ص26 / 33

معالجة اللغة البشرية الآليات والتقنيات والمبادئ

Human Language Processing: Mechanisms, Techniques, and Principles

أد،عزالدين عماري *	أد، صالح غيلوس
جامعة المسيلة (الجزائر)	جامعة المسيلة (الجزائر)
Azzedine.ammari@univ-msila.dz	Salah.ghilous@univ-msila.dz

معلومات المقال

تاريخ الارسال:
2025/11/20

الكلمات المفتاحية:

- ✓ ذكاء اصطناعي
- ✓ لغة بشرية
- ✓ معالجة اللغة الطبيعية
- ✓ فهم آلي
- ✓ التفاعل الإنساني الحاسوبي

الملخص:

يتناول هذا البحث **العلاقة المحورية والتكاملية** بين **الذكاء الاصطناعي (AI)** واللغة البشرية، مؤكداً على الأهمية القصوى لتقنيات **معالجة اللغات الطبيعية (NLP)** كجسر يُمكّن الأنظمة الحاسوبية من فهم وتحليل النصوص والكلام البشري على مستويات متعددة (نحوية، دلالية، وعملية). يستعرض البحث **التطور التاريخي** لتقنيات AIفي هذا المجال—من المنهجيات القائمة على القواعد إلى ثورة **التعلم العميق** ونماذج اللغة الكبيرة—**(LLMs)** موضحاً الأسس النظرية التي يقوم عليها الفهم الآلي. كما يسلط الضوء على **التطبيقات العملية** الجوهرية لـNLP ، مثل الترجمة الآلية، تحليل المحادثات واستخلاص المشاعر، واستخراج البيانات. ويختتم البحث بمناقشة **التحديات القائمة**)كالغموض اللغوي وتنوع اللهجات) ورسم **الفرص المستقبلية** لاستخدام الذكاء الاصطناعي لتطوير التفاعل الإنساني الآلي.

Article info

Abstract : (not more than 10 Lines)

This research addresses the pivotal and complementary relationship between Artificial Intelligence (AI) and human language, emphasizing the paramount importance of Natural Language Processing (NLP) techniques. NLP acts as a crucial bridge, enabling computer systems to understand and analyze human text and speech across multiple levels (syntactic, semantic, and pragmatic). The paper reviews the historical development of AI technologies in this domain—from rule-based methodologies to the revolution of Deep Learning and Large Language Models (LLMs)—clarifying the theoretical foundations upon which machine comprehension is built. Furthermore, it highlights the essential practical applications of NLP, such as machine translation, conversational analysis, sentiment extraction, and data retrieval. The research concludes by discussing existing challenges (like linguistic ambiguity and dialect variation) and outlining the future opportunities for leveraging AI to advance human-machine interaction.

Keywords:

- ✓ *Keyword*:
- ✓ الكلمة المفتاحية (عربي),الترجمة الإنجليزية *(English Translation)*
- ✓ *Artificial Intelligence (AI)*
- ✓ *,Human Language / Natural Language*
- ✓ *,Natural Language Processing (NLP)*

✓ *,Machine Understanding / Computational Comprehension* ✓ *Human-Computer Interaction (HCI) / Human-Machine Interaction*	

مقدمة:

تُمثل اللغة البشرية عبر تاريخها الطويل، أعمق وأعقد مظهر للذكاء والإدراك الإنساني؛ فهي الأداة التي صاغت الحضارات، ونقلت المعرفة، وشكلت الوعي ومع بزوغ فجر الذكاء الاصطناعي (Artificial Intelligence - AI)في منتصف القرن العشرين، تحول هذا المظهر الإنساني الجوهري إلى ساحة التحدي الأكبر للآلة :هل يمكن للحاسوب أن يفهم لغتنا ويتحدث بها؟

العلاقة بين هذين المجالين ليست مجرد علاقة تطبيقية؛ بل هي علاقة فلسفية وعلمية عميقة، تبحث في جوهر الفهم والمعرفة، فبقدر ما تنجح الآلة في معالجة اللغة البشرية، فإنها تقترب من محاكاة العمليات الذهنية المعقدة التي يقوم بها العقل البشري.

فالتفاعل اليوم أكثر أهمية من أي وقت مضى، مدفوعًا بالكم الهائل من البيانات النصية والرقمية المتاحة. هنا، تبرز معالجة اللغات الطبيعية (Natural Language Processing - NLP) كحجر الزاوية، الذي يربط بين الذكاء الاصطناعي واللسانيات، فمهمة NLP هي تزويد الآلة بالآليات والتقنيات اللازمة لفك شفرة اللغة البشرية الغامضة، وغير المهيكلة، وتحويلها إلى بيانات منظمة يمكن اتخاذ القرارات بناءً عليها.

لقد شهد العقد الأخير قفزات نوعية في هذا المجال، خاصة مع ظهور نماذج التعلم العميق وما يسمى بـ النماذج اللغوية الكبيرة(LLMs) ، التي أثبتت قدرة فائقة على فهم السياق المعقد، وتوليد نصوص إبداعية، وتقديم إجابات متماسكة، مما أعاد تعريف مفهوم التفاعل الإنساني الحاسوبي.

تستعرض هذه المقدمة الأهمية القصوى لتمكين الذكاء الاصطناعي من فهم اللغة البشرية، وتسلط الضوء على الأساس النظري والتقني الذي تستند إليه هذه المعالجة، وصولاً إلى استشراف التحديات والآفاق المستقبلية التي تنتظر هذا الحوار المستمر بين الإنسان والآلة.

- **الذكاء الاصطناعي واللغة الطبيعية :**

الذكاء الاصطناعي صار اليوم واحد من أهم الأدوات المساعدة للحاسوب تفهم لغتنا بشكل أفضل لأنظمة معالجة اللغة الطبيعية تعتمد على تقنيات متطورة تخلي الأجهزة تقدر تحلل النصوص وتفهمها بطريقة تشبه طريقة تفكير الإنسان. من بين هذه التقنيات،(التعلم الآلي والتعلم العميق)، التي تساعد الأنظمة ، وتكتشف أنماط اللغة وتستخلص المعاني من النصوص سواء كانت مكتوبة أو منطوقة.[1]

حيث يفتح المجال لتطبيقات كثيرة في مجالات متنوعة، مثل: (الترجمة الآلية، وتحليل المحادثات، واستخراج المعلومات من كميات كبيرة من النصوص)، مثل: برامج الترجمة الحديثة التي تعتمد على نماذج لغوية مبنية على الذكاء الاصطناعي، وهذا يجعلها تترجم النصوص بدقة أعلى من الطرق التقليدية؛ لأنها تراعي سياق الكلام والمعاني الدقيقة.

إن تقنيات معالجة اللغة الطبيعية تساعد في بناء أنظمة تقدر تحلل المشاعر الموجودة في النصوص، وتفهمها بشكل أعمق، حيث يُعد فهم العلاقة بين الذكاء الاصطناعي، واللغة الطبيعية أمرًا أساسيًا لتطوير تطبيقات عملية قادرة على تعزيز التفاعل بين الإنسان والحاسوب، وتحسين جودة التحليل اللغوي واستخراج المعلومات .[2]

- **الذكاء الاصطناعي لفهم اللغة البشرية:**

أصبحت الأنظمة الحاسوبية قادرة على تحليل النصوص والكلام بشكل شبه مشابه للفهم البشري. بدأت البحوث الأولى في هذا المجال باستخدام القواعد النحوية والصرفية المباشرة لتفسير الجمل، لكنها كانت محدودة في قدرتها على التعامل مع السياقات المعقدة والمعاني المتعددة.

مع ظهور تقنيات التعلم الآلي والتعلم العميق، أصبح بالإمكان تدريب النظم على كميات كبيرة من البيانات اللغوية، ما يمكّن الحواسيب من التعرف على الأنماط اللغوية ، واستخلاص المعاني ضمن سياقات متعددة.[3] وقد أدى هذا التطور إلى تحسين دقة النظم في التطبيقات العملية، مثل: الترجمة الآلية، وتحليل المحادثات، واستنتاج المعلومات من المستندات النصية .

كما أن تطور الشبكات العصبية الاصطناعية وموديلات اللغة الكبيرة (Large Language Models)ساهم في تحسين قدرة الحواسيب على فهم اللغة البشرية بشكل أكثر دقة ومرونة، حيث يمكن للنظام الآن التعامل مع التعابير الاصطلاحية، والفكاهة، والعبارات متعددة المعاني وقد أظهرت

الدراسات أن الدمج بين الذكاء الاصطناعي واللسانيات الحاسوبية يفتح آفاقًا واسعة لتطوير تطبيقات ذكية مثل المساعدين الرقميين والدردشة الآلية (Chatbots) التي تقدم استجابات متقاربة للطريقة التي يتفاعل بها البشر.[4]

المعالجة الآلية للغات الطبيعية – المبادئ والتقنيات: تعتبر المعالجة الآلية للغات الطبيعية واحدة من الفروع الرائدة في الذكاء الاصطناعي، وهي تهدف إلى تمكين الحواسيب من فهم اللغة البشرية وتحليلها وإنتاج نصوص جديدة بأسلوب طبيعي .تعتمد هذه المعالجة على مجموعة من المبادئ التقنية واللغوية التي تدمج بين علوم الحاسوب واللسانيات، مما يسمح للأنظمة بالتعرف على القواعد النحوية والصرفية، وفهم البنية الدلالية للنصوص . وتشمل أساليب معالجة اللغة الطبيعية الحديثة مستويات تحليل متعددة، منها:

- **تحليل الصرف والنحو**: يسعى هذا التحليل إلى تحديد العناصر الأساسية لكل من الكلمات والجمل، والتحقق من صحة البنية اللغوية، مما يساعد الحاسوب على فهم التراكيب اللغوية بشكل أفضل) .

- **تحليل المعاني والدلالات:**

يركز على استخراج المعاني والصلات بين الكلمات داخل النصوص، مع الأخذ بعين الاعتبار السياقات المتنوعة، مما يعزز من قدرة الحاسوب على فهم النصوص الطبيعية.

- **تحليل السياق واستخراج المعلومات:**

يساعد هذا النوع من التحليل الأنظمة على استخراج المعلومات الأساسية من النصوص الكبيرة والمعقدة، ويعتبر أساساً لتطوير تطبيقات عملية، مثل: أنظمة الترجمة الآلية والمساعدين الرقميين.[5] من خلال دمج هذه المبادئ والتقنيات، ويمكن للنظم الحاسوبية التعامل بذكاء مع النصوص، والقيام بمهام مثل : التعرف على الكلام، وتصنيف النصوص، وتحليل المشاعر، وتحليل البيانات اللغوية الكبيرة، و يساهم التحسين المتواصل في نماذج اللغة الكبيرة في رفع مستوى فهم اللغة الطبيعية، ويعزز من التفاعل بين الإنسان والحاسوب بشكل أكثر سلاسة.

المراحل الأساسية للمعالجة الرقمية للغة:

محاولة لتعليم الكمبيوتر مهارات "القراءة"، "الفهم"، و"الكلام "بأسلوب يشبه البشر عندما تقرأ جملة ما، يقوم عقلك بعدد كبير من العمليات الذهنية، بشكل سريع لفهمها؛ لأن المعالجة الرقمية للغة تعمل على تحليل هذه العمليات وتجزئتها إلى خطوات منهجية قابلة للتنفيذ، بواسطة الكمبيوتر.

- **المرحلة الأولى**:

- **التنظيف والتحضير ما قبل القراءة :**

تمثل هذه المرحلة إعداد النص الخام وتنظيمه، تمامًا مثل : تجهيز المكونات قبل البدء في عمل ما.

- **التجزئة و تقسيم النص :**

يتم تقسيم الجمل إلى كلمات فردية .مثال" :الكمبيوتر يفهم اللغة" تتحول إلى "الكمبيوتر"، "يفهم"، "اللغة وذلك بإزالة الكلمات الزائدة :(Stop Words) تحذف الكلمات التي تحمل قيمة محدودة في المعنى، أو التي تُستخدم غالبًا كروابط دون تأثير كبير على السياق، مثل: (من، على، الـ، في .لأجل التركيز على الكلمات ذات الأهمية الجوهرية ، واختزال الكلمات إلى الجذر (Stemming/Lemmatization).إذ ترجع جميع الكلمات المختلفة أو مشتقاتها إلى أصلها.مثل: الكلمات "يكتبون"، "كاتب"، و"كتابة" تُصبح "كتب ."

- **المرحلة الثانية**:

فهم التركيب والمعنى القراءة والتحليل، يبدأ الكمبيوتر في تحليل هيكل الجملة وفهم معناها ، بتعريف دور الكلمات (POS Tagging) ، فيتم تصنيف كل كلمة بناءً على وظيفتها النحوية في الجملة : هل هي اسم؟ أو فعل؟ أوصفة؟ والهدف هو فهم الدور القواعدي لكل عنصر في النص من خلال.

- **تحليل بنية الجملة** :(Parsing) يتم إنشاء مخطط يعرض العلاقات بين الكلمات داخل الجملة لفهم الهيكل العام، مثل تحديد الفاعل والمفعول به.

- **تحديد المعنى المقصود**:(Word Sense Disambiguation) :

يساعد الكمبيوتر على اختيار المعنى الصحيح للكلمة استنادًا إلى سياق استخدامها .مثل: كلمة "عين" في جملة "شربتُ من العين" تعني ينبوع ماء، بينما في "رأيت عين الطائر "تشير إلى عضو الرؤية .

- **التعرف على الكيانات المهمة** (NER - Named Entity Recognition) :

يتم تحديد وتصنيف الأسماء الخاصة التي تشير إلى أشخاص، وأماكن، ومنظمات، أو أزمنة معينة ضمن النص، كـ: "زار أحمد المسيلة أمس"، يتم التعرف على" أحمد" كاسم شخص، "المسيلة" كموقع، و"أمس" كوقت .هذه المراحل تمثل الأساسيات التي يعتمد عليها الذكاء الاصطناعي، لفهم ومعالجة النصوص بأسلوب يحاكي طريقة تفكير الإنسان .

- **التعلم العميق والقفزة نحوLLMs :**
- **مفهوم التعلم العميق** (Deep Learning) .
- **التعلم العميق:** مجموعة فرعية من التعلم الآلي، تستخدم الشبكات العصبية الاصطناعية (Artificial Neural Networks). ذات الطبقات المتعددة (الشبكات العميقة) لتحليل البيانات بشكل مشابه للدماغ البشري.
- **الآليـــــــة :**

تتميز هذه الشبكات بقدرتها على استخلاص وتمييز الميزات (Features) بشكل تلقائي من البيانات الأولية، (كالصور، أو النصوص). فبدلًا من أن يحدد المبرمج الميزات التي يجب البحث عنها (مثل الحواف أو الزوايا في صورة)، تقوم الشبكة العصبية العميقة بتعلم هذه الميزات بنفسها عبر طبقات متعددة من المعالجة.

- **الأهمية في اللغة :**

في سياق اللغة، سمحت هذه الشبكات بمعالجة السياق (Context) والفروق الدقيقة في القواعد والدلالات بشكل أفضل بكثير من النماذج الإحصائية التقليدية.

- **ولادة نماذج المحولات**.(Transformers) :

كانت النقطة الفاصلة في تطور التعلم العميق ونماذج اللغة هي ظهور بنية المُحولات (Transformer Architecture)في عام 2017.

- آلية الانتباه :(Attention Mechanism) :

تقدم المحولات آلية الانتباه التي سمحت للنموذج بـ "ترجيح" أهمية الكلمات المختلفة داخل الجملة ،أو النص عند معالجة كلمة معينة. هذا يمكّن النماذج من التقاط التبعيات طويلة المدى(-Long Range Dependencies) بين الكلمات حتى لو كانت متباعدة جداً في النص.

- **التدريب المتوازي :**سمحت هذه البنية بتدريب النماذج بكفاءة عالية على أجهزة قوية، مما فتح الباب أمام بناء نماذج أكبر بكثير.
- **النماذج اللغوية الكبيرة :(LLMs)** المفهوم والتوسعة.
- **نماذج اللغة الكبيرة** (Large Language Models) :هي نماذج تعتمد على التعلم العميق وبنية المحولات، وتتميز بثلاثة خصائص أساسية جعلتها ثورية:
- **الحجم الهائل**:(Scale)
- **المعلمات :(Parameters)** يُقصد بـ "الكبيرة" حجم الشبكة العصبية، حيث تحتوي هذه النماذج على مليارات المعلمات (أوزان وتحيزات يتم تعديلها أثناء التدريب). كلما زاد عدد المعلمات زادت قدرة النموذج على تخزين وتعلم الأنماط المعقدة في اللغة.
- **البيانات :**يتم تدريبها على كميات هائلة (Terabytes) من البيانات النصية المأخوذة من الإنترنت، الكتب، والمقالات، مما يمنحها قاعدة معرفية واسعة جداً.
- **التدريب المُسبق (Pre-training) والتعلم المنقول:(Transfer Learning)**
- **التدريب العام :** يتم تدريب النماذج أولاً على مهمة عامة جدًا، وهي التنبؤ بالكلمة التالية في النص أو ملء الفراغات من خلال هذه المهمة البسيطة المتكررة، وتتعلم النماذج الهيكل الكامل للغة (القواعد، الدلالات، السياقات).
- **الضبط الدقيق** :(Fine-tuning) بعد التدريب العام، يمكن "ضبط" النموذج على مجموعة بيانات أصغر ومحددة لتأدية مهمة معينة بدقة، مثل: الإجابة عن الأسئلة أو تحليل المشاعر.
- **الذكاء متعدد المهام** :(General-Purpose AI) .

على عكس نماذج NLP القديمة التي كانت تخصص لمهمة واحدة ،مثل :(الترجمة أو التصنيف)، أصبحت LLMs نماذج متعددة الأغراض .(General-Purpose) ،يمكنها أداء مجموعة واسعة من المهام اللغوية دون الحاجة لإعادة تدريبها بالكامل، لكل مهمة فقط عن طريق تغيير الإرشادات المُقدمة لها.(Prompting) .6

- **خاتمة:**

- يبقى التحدي الأكبر للذكاء الاصطناعي مرتبطًا بجوهر الإنسانية وهو اللغة البشرية .فعلى مر العصور، مثلت اللغة أعمق وأعقد تعبير عن الذكاء والإدراك البشري؛ كونها الأداة التي بُنيت بها الحضارات، ونُقلت المعرفة، وشُكلت بها الوعي والفكر .ومع انطلاقة عصر الذكاء الاصطناعي(AI) ، أصبحت اللغة هذا العنصر الإنساني الفريد ميدانًا رئيسيًا لاختبار قدرات الآلة من خلال تقنيات معالجة اللغات الطبيعية(NLP) ، وخصوصًا النماذج اللغوية الكبيرة(LLMs) ، التي غيرت بشكل كبير النظرة التقليدية إلى قدرة الحاسوب على فهم اللغة البشرية واستخدامها. فلم يعد التساؤل عن إمكانية الحاسوب لفهم لغتنا، واستعمالها محصورًا في نطاق النفي المطلق، إذ أظهرت هذه النماذج قدرات مذهلة في محاكاة الفهم، والابتكار اللغوي. ورغم ذلك، يبقى المجال مفتوحًا لمزيد من التطوير والتقدم نحو تحقيق إدراك آلي حقيقي، يتجاوز مجرد التحليل الإحصائي للكلمات ليصل إلى فهم أعمق للسياق، والقصد، والنوايا الإنسانية.

الهوامش :

[1] Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (الطبعة الثالثة). Youcanprint.

[2] - رضا بابا أحمد: اللسانيات الحاسوبية مشكل المصطلح والترجمة مخبر المعالجة الآلية للّغة العربية، الجزائر، دط، دت، ص14.

[3] -Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

[4]- نفسه، ص 55

[5]- الدراوي، ص. 60

[6] **-Devlin, J., et al. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding."** *NAAC*