الفصل

الخامس

حدودُ
الذكاء الاصطناعي

ويشتمل هذا الفصل على النقاط التالية:

- ✓ أولا: الروبوتات
- ✓ ثانيا: رؤية الكمبيوتر
- ✓ ثالثا: التعرف على الكلام

الفصل الخامس

حدودُ الذكاءِ الاصطناعيّ

ما هي المجالاتُ الرئيسيةُ للبحثِ والتطويرِ في الذكاءِ الاصطناعيِّ؟

ينقسمُ العملُ في الذكاءِ الاصطناعيِّ عمومًا إلى عددٍ من الحقولِ الفرعيةِ التي تعالجُ المشاكل العمليةَ الشائعةَ، وإن كانت صعبةً، أو تتطلب أدواتٍ أو مهاراتٍ مختلفةً.

بعضٌ منها وأكثرها بروزًا هي الروبوتات، والرؤية الحاسوبية، والتعرف علي الكلام، ومعالجه اللغة الطبيعية.

وفيما يلي شرح موجز لكل منها:

أولا: الروبوتات:

ينبغي أن تتطلبَ الروبوتاتُ وصفًا قليلًا، ينطوي علي بناء الأجهزةِ القادرةِ علي أداء المهامِ الفعليةِ، وهي تماثلُ التفكيرَ الإنسانيَ، ومحاكاةَ شكلِ الإنسانِ، ولكن بالطبع هذا ليس ضروريًا، والكثير من العمل الجاري يسعى إلى تطوير ربوت أخف وزنًا، وأكثر مرونةً، وأقوي في المادةً وفي أساليب السيطرة، فضلًا عن التصاميم التي غالبا ما تكون مستوحاةً من الطبيعة.

لكن ما يميز حقًا الأبحاثَ الروبوتيةَ في الذكاء الاصطناعي هو أنها من أكثر الآلات الميكانيكية أتمتةً (ميكنة)، كما أنها تُعدُ محاولة لبناء أجهزةً قادرةً علي المزيد من إنجاز المهام العديدة، فعلي سبيل المثال، يوجد كثير من الأنواع من الآلات ذات الغرض المخصوص مثل تلك الخاصة بِحزمةِ أغذيةٍ محددةٍ، أو تعليب المنتجات في علب الشحن، أو في الحاويات، ولكن صُنْعُ جهازٍ واحدٍ قادرٍ علي التعامل مع مجموعةٍ واسعةٍ من الأشكالِ أمر صعب، ولا تزال الأحجامُ والأوزانُ الخفيفة تشكل تحديًا في أبحاث الذكاء الاصطناعي، وهي مسألة مهمة في التكيف مع التغيرات الآلية.

فالبيئات البحثية مستمرة في العمل على تطوير روبوتات قادرة على صناعة سيارات مستقلة، تقطع الطرق وتنجز المسافات في سباق يُقام مع السيارات والدراجات التي سيطر عليها الإنسانُ.

وعلي الرغم من كل ما يصاحبُ ذلك من الجدةِ والتنبؤِ، فقد أصبحت تكنولوجيا الذكاء الاصطناعي تفتح آفاقًا جديدة كاملة من الفرص الاقتصادية من خلال تمكين الروبوتات للعمل بالمهام التي لا يستطيع الناس إنجازها، كما أنها ذات قيمةٍ كبيرةٍ في إنجاز كل أنواع المهام الخطيرة أو التي تمثل ضررً على الناس المقيمين بها. من هذه المهام التعدين والزراعة والغوص في قاع البحر، والقضاء علي الآفات الزراعية من خلال استهدافها مع المفترسات الميكانيكية الخاصة بالحشرات، أو تنظيف الحوادث الصناعية.

ومن بين هذه المجالات الواضحة أيضًا استكشاف الفضاء، ففي عام 1993م ، أرسلت وكالة مكوك فضاء مع سبعة أشخاص علي متن سفينة فضائية في مهمة لإصلاح محطة الفضاء هابل؛ وكان الهدف من ذلك هو القيام بعملية دقيقة بشكل استثنائي لتصحيح تلسكوب بصري خاص بالفضاء، وكانت هذه الصيانة الأولى من أصل خمس صيانات خاصة بمحطة هابل (2) المأهولة بالبعثات في 2004، ونظرتْ بجديةٍ في استخدام اثنين من الروبوتات المسلحة الكندية "دعا" و"دكستر" بدلًا من رواد الفضاء للمهمة النهائية، ولكن كان محكومًا عليها أن تكون محفوفةً بالمخاطر نظرًا لغاية الدولة من الفن في ذلك الوقت.

ومع ذلك، من المرجح أن تكون أجهزةُ الروبوتات عمليةً لكثير من أنواع المهام ومن المرجح أن تقوم بأعمال في باطن الأرض في المستقبل القريب، مثل تحليل العينات الجيولوجية، والبحث عن الحياة البيولوجية، والتعدين على الكويكبات، وتحويل الأجسام الفلكية التي تهدد مسارات الأرض.

فرصة وكالة ناسا في المريخ من الأمثلةِ الرئيسية على هذا النهج، علي الرغم من أن درجة اعتمادها على تكنولوجيا الذكاء الاصطناعي أقل وضوحًا، وتكاد تكون أقرب إلى المستخدمة في المنزل، ولكن التحدي الحديث للروبوتات من وراء الدافع تمثل في صعوبة الحصول علي الموظفين البشريين في محطات "فوكوشيما النووية" بعد الانهيار.

وتعددت النشاطات للروبوتات فأصبحت تتنافس في تنفيذ مجموعة متنوعة من المهام المشتركة مثل قيادة سيارة المساعدة، وفتح الباب، وتحديد موقع وإغلاق صمام الأمان، وربط خرطوم النار في ماسورة.

ومن المجالات الأخرى في أبحاث الروبوتات هو العناية بالمسنين، وكان هذا بغرض النجاة من حالات الشيخوخة بدافع من الاقتصاد الديموغرافي في العديد من المجتمعات الغربية، ولا سيما في اليابان.

وهناك عدد من الجهود المبذولة في إطار طريقة لتقديم المساعدة الروبوتية لإعلام المسنين، والعاجزين، ولكن العميلة الأكثر تركيزًا تلك التي تركز علي مهام محددةٍ، مثل التأكد من أن المرضي يأخذون أدويتهم أو يساعدونهم علي التحرك من السرير إلى الكرسي المتحرك، علي الرغم مما تراه في الأفلام من الروبوت المنزلية مثل الروبوت "فرانك"، هي وسيله طويلة جدًا من توفير هذا النوع من المساعدة العامة في الرعاية البشرية.

وهناك فئة منفصلة من الروبوتات تقدم المساعدة النفسية بدلًا من الراحة الجسدية، علي سبيل المثال، الروبوت العلاجيُّ "بارو" يوفر فوائد "العلاج الحيواني" بالإضافة إلى معرفة المرض.

وكان ظهور "بارو" لتحسين التنشئة الاجتماعية، وزيادة الاسترخاء، وتعزيز الدافع. ومع ذلك، فإن هذه الروبوتات "العاطفية" مصطنعة بدون جدال. على يد البروفيسور "شيري" في معهد ماساتشوستس للتكنولوجيا الذي يقوم بدراسة الآثار الاجتماعية للتكنولوجيا، ويحذر من أن الأجهزة الميكانيكية التي تشجع على الترابط العاطفي بقوله إنها خادعة بطبيعتها، ويحتمل أن تكون ضارة للعلاقات البشرية.

أيضًا هناك روبوتات للترفيه، وهذه عادة تتخذ أشكالًا مجسمةً، مثل الرسومات الرقمية من قبيل المشتركة في الحدائق الموضوعة في مدينة ألعاب "ديزني لاند"، ولكنها أكثر مرونةً وتفاعليةً منها "بيبر" من الروبوتات، وموبايل"سوفت بنك"، الخاص بقراءة النوايا والاستجابة بشكل مناسب، وهو يستخدم حاليًا في استقبال الزوار لمتاجر "سوفت بنك" في اليابان، حيث يمكنه الإجابة على أسئلة محدودة حول المنتجات والخدمات، ولكن القيمة الأساسية هي لاشتراك وفرحة العملاء.

وكانت هناك أيضًا أجيال عديدة من الروبوتات في مثل اللعبة التفاعلية من hasbro سوني الروبوتية "الكلبة أيبو" وتهدف هذه الأدوات لبهجة الأطفال وسحر الكبار مع زيادة التطور والاستجابة التفاعلية بين الإنسان والآلة.

الخادمةُ الشخصية الميكانيكية حلمٌ قديمٌ مثل الروبوتات نفسها، ولكن الصورة الشعبية للخادمة الروبوتية، في نمط "روزي" علي التليفزيون في الرسوم المتحركة الكلاسيكية التي تظهر "jettons"، وهو لا يزال حلمًا بعيدًا.

أيضًا الحالة الفعلية للفن هو مثال من المكنسة الكهربائية المستقلة من روبوت "rumba"، والتي تنطلق حول الستائر، والسجاد، وتتجنب الخطوات، وتقوم بتجهيز نفسها للشحن في محطة الشحن الخاصة بها، كما أنها تحاول الابتعاد عن طريقك أثناء القيام بعملها.

الإصدار الأخير هو "rumba 980"، الذي يمكنه رسم خريطة قويمة لبناء منزلك، وتتأكد من أنه يقوم بتنظيف شامل للفضاء الخاص بك، في حين أن الإصدارات السابقة تحولت إلى آلات هشّة، في شكل عشوائي.

اثنتا عشرة مرحلة من التطورات الأخيرة الأكثر إثارة في هذا المجال، ويعرف باسم "سرب الروبوتات"، وهي مجموعات كبيرة نسبيًا مبرمجة على روبوتات بسيطة متفقة مع القواعد، وعندما يتم تطبيق هذه القواعد بالكامل علي مجموعات الفريق بأكملها، تُظهر الروبوتات سلوكًا معقدًا يُسمى "السلوك الناشئ"، وقد لوحظ هذا التأثير نفسه في عش النمل والأعضاء الموجوة في خلايا النحل المجتمعة لحل المشاكل، التي هي أبعد بكثير من فهم قدرات أي فرد، في حين أن "سرب الروبوتات" يمكن أن يكون أي حجم، وتركز الكثير من البحوث على المقاييس الصغيرة مثل (الحشرات) وهو ما يعرف بـ "النانو الروبوتية".

يمكن أن تعمل مجموعات هذه الأجهزة معًا لتنفيذ بعض المهمات، على سبيل المثال، تحديد موقع الناس المحاصرين في انهيار المباني أو الكشف عن تسربات سامة، وينتهي التنسيق عن طريق تشكيل شبكات مخصصة، أو التواصل بين الأقران مع الوحدات القريبة. إلا أنه من الصعب التحديد الدقيق للفوائد والمخاطر المحتملة من هذه التكنولوجيا.

على الجانب الإيجابي أيضًا يمكن أن يحقق تقدمًا طبيًا هائلًا، مثل الأداء الناجح والسليم للعمليات الجراحية داخل الجسم.

تخيل مجموعة كاملة من الروبوتات في حجم الخلايا التي تحاكي وظيفة الجهاز المناعي، قادرة على البحث عن السرطانات التي تنتقل عن طريق الدم ومهاجمة الأجهزة داخل الجسم، أو صندوق مليء بآلات الروبوت بحجم الصراصير التي تسرع حول جمع الغبار من الغُرَفِ والجدران، وحشْوها في حقيبة صغيرة لسهولة التخلص منها.

وانطلاق الآلاف من الروبوتات في حجم حيوان الخُلْد لاستكشاف المعادن تحت الأرض، تليها عمال المناجم الآلية الصغيرة، ولكن هناك أيضا مخاطر كبيرة، ويمكن استخدام نفس التكنولوجيا التي قد تعالج السرطانات التي تنتقل عن طريق الدم لقتلك، أو ربما حتى للسيطرة عليك، أكثر الناس تعرف أن القضاء على النمل من المطبخ أمرٌ صعبٌ، فليس من السهل غزو جيش صغير منظم.

إن قدرة التطبيقات العسكرية أو الإرهابية من "سرب الروبوتات" هي حقًا تفكير فظيع جدًا، حيث إن الأبحاث ذات الصلة حول مجموعات متعاونة من الروبوتات المتعددة، في جداول أكبر، تهدف إلى تنسيق نشاط مجموعات من الروبوتات ديناميكيًا، عادةً من مورد حوسبة مركزي، على سبيل المثال "Kiva Systems"، ضمن إدارة مستودعات شركة الروبوتات التي تم شراؤها من قبل شركة أمازون في عام 2012، وتتمكن أفعال الروبوتات من جلب المنتجات من الرفوف إلى التعبئة، وكما أنها أفعال (البشرية)، ومن أجل إلهام وتشجيع البحوث على أنظمة متعددة الروبوتات، يتم تنظيم مسابقة "Robocop" سنويًا، حيث تتنافس الفرق للفوز بمسابقة كرة قدم الروبوتية، ويكون الاسم الرسمي هو "كأس العالم لكرة القدم الروبوت".

التطبيقات العسكرية كثيرة جدًا، وفي بعض الأحيان تكون خطرة، بينما يستحضر الخيال الشعبي حتى يتمكن من رؤية مجموعة من الجنود الروبوتية على غرار المنهي عنه من أن يركض على مسرح المعركة وهو يحمل السلاح، والحقيقة مختلفة جدًا، كما أنه لن يتم تصميم الروبوتات العسكرية لاستخدام الأسلحة، ومن أمثلة هذه الأسلحة البنادق التي يمكنها تحديد الأهداف وإطلاق النار بشكل مستقل، كذلك الطائرات من دون طيار التي يمكنها أن تقوم بنقل شحنات متفجرة إلى مواقع محددة، والألغام الأرضية التي تنفجر فقط عندما تكون سيارات العدو محددة داخل النطاق، وهذه الاحتمالات مزعجة للغاية بحيث تبذل جهود كبيرة من قِبل الأمم المتحدة.

والدراسة في المؤسسة العسكرية لها أخلاقيات فعالة في استخدام مثل هذه الذخائر الدقيقة لدعم أو استبدال الأفراد في مناطق الحرب المتعددة.

الإجماع الحالي هو كمسألة حذر، كيف يجب على الإنسان أن يكون "في حلقة " لجميع قرارات الاستهداف قبل سحب الزناد، ولكن ليس من الواضح تمامًا أن هذا أمر عملي، أو يمكن الدفاع عنه أخلاقيًا، نظرًا لأن طلب هذه المراجعة قد يعرض حياة الأشخاص للخطر على النقيض من بعض التطبيقات الأخرى الأكثر وضوحًا.

وفي ظلال تلك الروبوتات من مثل الأجهزة البسيطة التي تنفذ إجراءات روتينية (كما هو شائع في المصانع) للأنظمة المعقدة التي تفيد بيئتهم، وعقلهم، واتخاذ إجراءات تخصهم، وضبط خططهم ردًا على الملاحظات الجديدة، وبالتالي فإن حدود "فيلد" بعيدة كل البعد عن الوضوح، ولكن من المفيد وضع ذلك في الاعتبار متأخرًا عن التصور العام بشكل كبير، ومن السهل لتصوير فيديو من روبوت جذاب مع عيون كبيرة ووجه تعبيري يتفاعل بطرق مناسبة اجتماعيًا مع متدرب مدَرَّبٍ، ولكن بالنسبة للجزء الأكبر هذه النظم فهو أكثر هشاشة بكثير مما يتوقعه الناس، على الأقل حتى الآن.

ثانيا: رؤية الكمبيوتر:

رؤية الكمبيوتر تركز بشكل أساسي على تجهيز أجهزة الكمبيوتر مع القدرة على "الرؤية"، بمعنى تفسير الصور المرئية، والعمل في مجال الرؤية الحاسوبية يوازي الانتقال من الأنظمة الرمزية إلى تعلم الماكينة.

وقد ركزت الجهود المبكرة على صياغة الخوارزميات المستخدمة المعروفة المتحققة من الصور البصرية وأوصاف الأشياء المثيرة للاهتمام؛ للبحث عن عناصر ذات معنى دلالي مثل الخطوط والمناطق وما إلى ذلك، والتي غالبًا ما يتم تجميعها بعد ذلك في كيانات أكبر وأكثر عمومية. على سبيل المثال، هناك برنامج مصمم لتحديد الكرسي يتمكن من البحث عن الأرجل، المقعد، الظهر، وما شابه ذلك. لكن النهج الأكثر حداثة هو استخدام الآلة للتعلم، وغالبًا ما تكون أنواع متخصصة من الشبكات العصبية (تسمى الشبكات العصبية التلافيفية، أو شبكات CNN)، لبناء نماذج من الأشياء من مجموعات كبيرة من الأمثلة.

"سي إن" تتحدث بشكل جيد جدًا، أثناء البحث عن أنماط في أقسام صغيرة متداخلة من صورة ما، ثم يمكن نشر ما يتعلمون بداية إلى الأقسام المجاورة ثم إلى مناطق أكبر بشكل تدريجي للصورة، باستخدام هذه التقنيات.

كان التقدم الأخير في "فيلد" سريعًا جدًا، على سبيل المثال، الدقة في صورة شبكة محددة "ImageNet Frontiers" السنوي لمركز (صناعة المخابرات) Artifcial Intelligence 55

وهو تحدٍّ بصريٍّ واسع النطاق، هدفه هو الكشف عن مائتي نوع من الكائنات وتوطينها في 150000 صورة تحتوي على ألف فئة للكائن، زادت بشكل كبير معدلات الخطأ في مجموعة من خمسة في المئة، بانخفاض من عدة مرات هذا فقط قبل عدة سنوات، المسابقة الثامنة عشر تتوسع الآن لتمييز الأشياء في الفيديو وللأوصاف الأكثر مشاهدة للرواية، مثل "ركل الصبي الكرة لكنه أخطأ الهدف".

لكن الوعد بهذا "الفلتر المصور" يمتد إلى أبعد من مجرد الرؤية البصرية، وله طريقة مختلفة للتفكير في رؤية الكمبيوتر، أو المعالجة البصرية بشكل عام، هي أنها تأخذ مدخلات لصورٍ ثنائية الأبعاد تمثل الضوء المنبعث من الأسطح ثلاثية الأبعاد، ثم تفسر أو تعيد بناء نموذج من المشهد الأصلي، الذي قد يعيد بناء المشهد على أساس جيد، على سبيل المثال، صور متعددة من وجهات نظر مختلفة، ومعرفة الهندسة والفيزياء للضوء، وإعادة تحديد الأسطح المختلفة، وفهم خصائص الأشياء في العالم الحقيقي (عادةً ما يركب الناس خيولًا، ليس العكس).

إن العالم الحقيقي ثلاثيُّ الأبعاد يطيع قواعد معينة من تكوين وتقييد هذه القواعد في العرض البسيط ثنائي الأبعاد المتوقع لعين للإنسان، أو كاميرا رقمية. (هذه هي القواعد التي تنتهكها الأوهام البصرية). ومع ذلك، فإن الأساليب نفسها لديها الكثير من التطبيق الأوسع، في حين أن عيوننا ومعظم عينات الكاميرات مختلفة في الضوء المعاد، هناك جميع أنواع أجهزة الاستشعار التي تجمع البيانات حول العالم الحقيقي بخلاف ما يمكن للبشر رؤيته، يمكن للأجهزة الخاصة، على سبيل المثال، قياس الأشعة تحت الحمراء (أثناء درجة الحرارة)، وإعادة تعديلها في الإشارات (على سبيل المثال، الرادار والاهتزازات). نفس القواعد الأساسية والتقنيات التي تستخدم لمعالجة الضوء، يمكن تكييفها بشكل مناسب، ويتم تطبيقها على تفسير وإعادة بناء المشاهد على أساس هذه الإشارات غير المرئية.

هناك "مشاهد" تطيع بعض القيود المادية والقواسم المشتركة، ولكن لا يمكن رؤيتها من حيث المبدأ على الإطلاق (على الرغم من استخدامها الأدوات القائمة على الكمبيوتر التي يمكننا استخراج تصور لها)، في مثل المستخدمة في موقع وتشكيلات النفط تحت

الأرض، واستكشاف أورام الدماغ، وكشف العيوب في السدود الخرسانية تحت الضغط. طالما لدينا معرفة واعية بشأن المادة وخصائص المجال الذي نقوم بفحصه، ولدينا بعض الطرق لجمع الإشارات التي تعرض هذه النطاقات في الصور وبطرق نفهمها، إذن يمكننا استخدام تقنيات رؤية الكمبيوتر، والتي يتم تفسيرها على نطاق واسع، لمعالجتها.

ومن حيث المبدأ، لا يجب أن تكون المشاهد ولا الصور مادية، ما دام تطيع المجالات قواعد معينة، وتمثل الصور مجموعة من نقاط البيانات ذات الأبعاد الدنيا المعروفة بمطابقة عناصر المجال، يمكن معالجة البيانات لتوفير نظرة ثاقبة في بنية المجال، بمعنى آخر، يمكن لأجهزة الكمبيوتر "رؤية" الأشياء التي لا يمكننا القيام بها.

نفس الشيء ينطبق على الكثير من الحيوانات، وهذا ليس غموضًا كما يبدو، فعلى سبيل المثال، ترى الخفافيش باستخدام أصوات معادة، والطيور معظمها معتاد على رؤية الألوان التي لا يستطيع البشر رؤيتها، فهي مهارة لديهم تُستخدم لاختيار الأصحاب، وإشارة الجوع.

التطبيقات الرئيسية لتقنية رؤية الكمبيوتر:

يعتمد عدد لا حصر له من مشكلات العالم الحقيقي على تحديد أماكن الاهتمام وتحديد موقعها في بيئة معينة، يوجد منها مهام بسيطة على ما يبدو، مثل تأرجح مطرقة في الأظافر، أو تراص الأطباق، أو رسم البيوت، أو مروج القص، واختيار تعتمد فيه الثمار الناضجة على معرفة مكان الأشياء، وكانت التقنية اللازمة للتصرف وفقًا لهذه المعلومات هي الهندسة الميكانيكية الأساسية والروبوتات، وهي متاحة لبعض الوقت، ولكنها عندما كانت متوفرة كانت تقتصر على البيئات؛ حيث كانت الأشياء ذات الاهتمام في مواقف مسبقة، مغمورة، مثل على أوعية المصنع، لكن التطورات الحديثة في رؤية الكمبيوتر تجعل من الممكن أداء المهام المادية مثل هذه في بيئات العالم الحقيقي أقل تنظيمًا، وعلى مدى العقود القليلة القادمة، نحن على الأرجح سنشهد توسعًا دراماتيكيًا في فئات المهام، ولتلك الوظائف ما يمكن تنفيذها بواسطة الآلات.

المجال الرئيسي الثاني للتطبيق هو المعلومات نفسها: لقد انتهينا إلى حد كبير من الانتقال من الطرق المادية المستندة إلى الورق لالتقاط المعلومات وإبلاغها (النص والرسوم البيانية والصور وما إلى ذلك) لإدارة البيانات في شكل إلكتروني، لكن البيانات التي نجمعها ونخزّنها ونشاركها في قواعد المعلومات الاستخبارية في صناعة المخابرات "Artifcial Intelligence 57"

أصبح مرئيًا بشكل متزايد في الطبيعة تطور للكاميرا الرقمية، خصوصًا عند تم دمجها في أجهزة الاتصالات في كل مكان مثل الهواتف الذكية، فقد خفضت تكلفة التقاط الصور ومشاركتها إلى الصفر تقريبًا، بدلاً من النقر على لوحة مفاتيح صغيرة، فتجد الآن الكثير من الناس يتعامل ببساطة مع زر "انقر" ويتم إرسال الصورة. ونتيجة لذلك، فإن نسبة المعلومات البصرية، من خلال الإنترنت قد تضخمت، فتجد أن وحدة فيديو واحدة من المتوقع أن تشكل 84٪ من جميع متاجر الإنترنت.

وفقًا لدراسة حديثة للصناعة المشكلة هي أنه على عكس البيانات النصية، التي يمكننا تفسيرها إلكترونياً لأغراض الفهرسة والاسترجاع، فإنه ليس لدينا طريقة لإدارة الصور ومقاطع الفيديو ما لم تكن تعرض على المصدر أو يصنفها الإنسان، فأنت قد تُفاجأ عندما تعلم أنه عند إجراء بحث "Google" بالنسبة إلى الصور، فأنت لا تبحث في الواقع عن الصور نفسها، بل عن التصنيفات والنصوص المصاحبة التي تقترح ما قد يظهر، فهذا هو السبب في أن عمليات البحث هذه أقل بكثير من استرجاع صفحة الويب. لذلك فإن الجزء الأكبر من العملية الإلكترونية يتمثل في نقل البيانات من النماذج النصية إلى البصرية.

ونحن لدينا توسع في المعلومات من خلال الشبكات الرقمية، إلا أن تقنيات رؤية الكمبيوتر تقدم الوعد بإدارة كل هذا تلقائيًا، مثل برامج التعرف على الوجه فهي بالفعل تستخدم لأغراض متنوعة مثل الأمن القومي من خلال زيادة صور أصدقائك على الفيسبوك.

ولكن سرعان ما ستتسع القدرة على تفسير وتوسيم الصور لتشمل كل الصور تقريبًا،

وعموماً فإن تكنولوجيا الرؤية الحاسوبية قد تصل في الوقت المناسب لمساعدتنا من الغرق في محيط من المعلومات الخاصة.

ثالثا: التعرف على الكلام:

كان استخدام الكمبيوتر للغة بطريقة الذكاء الاصطناعي على النقيض من البشر، الذين من المفترض أن تتحدث عنهم الكتب قبل الكمبيوتر.

يعتبر التعرف على الكلام أكثر تشويشًا بشكل كبير من معالجة اللغة المكتوبة في جزء كبير بسبب التقلب والضوضاء الكامنة في تيارات الصوت من اللغة المنطوقة حيث أن فصل "الإشارة" عن "الضوضاء"، ونقلها إلى الكلمات المكتوبة المناسبة، مهمة شاقة للبشر وكذلك أجهزة الكمبيوتر، كما يمكن لأي مستهلك من "الطبقات الدنيا" المشاهدة لشاشة التلفزيون.

لكن بداية المشكلة تتمثل في فصل الكلمات من أصوات الخلفية، مثل ما اكتشفه الباحثون الأوائل في هذا المجال بسرعة، فليس هناك اختلاف واضح بين الكلمات، خلافًا لما قد تفكر به عندما تستمع لشخص يتحدث في معنى كبير وهو منقول أيضا عن طريق كيفية تغيير الحجم والنبرة وهو ما يسمي "علم العروض أو الموسيقى" من قبل اللغويين، ففي اللغة الإنجليزية مثلًا، يمكنك تغيير معنى الكلام تمامًا عن طريق زيادة درجة الصوت في النهاية.

مع كل هذه التحديات، إنها معجزة حقًا لأن المشكلة كان لا يمكن حلُّها على الإطلاق، فقد حاولت معظم الجهود التعرف على الكلام في وقت مبكر لتبسيط المهمة عن طريق الحد من المفردات، التي تعمل في مجال مبسط (مثل لعب الشطرنج)، والتي تتطلب المتحدث للتوقف مؤقتًا بين الكلمات، وأيًا من التصميم هو متحدث معين ويتطلب جلسات تدريبية مكثفة (لكل من المتكلم البشري والجهاز).

في محاولة للقفز والتقدم في هذا المجال، في عام 1971 قامت DARPA بتمويل مسابقة لمدة عام من أجل الكلام المستمر (بمعنى عدم وجود توقف مؤقت بين الكلمات)،

باستخدام مفردات لا تقل عن ألف كلمة، سواء كان أي من المتسابقين قد نجح، ومن هنا كانت مسألة الجدل، ورفضت الوكالة تجديد التمويل بعد هذه الفترة الأولية، حتى أنه أعاد إحياء اهتمامه بعد حوالي عشر سنوات في عام 1984، بينما استخدمت الفِرَقُ في هذه المسابقة مجموعة متنوعة من التقنيات المختلفة، يمكن وصف معظمها تقريبًا على أنها محاولات للتدوين، ويجلب الحكمة المقبولة من مجموعة متنوعة من الحقول، مثل بناء الجملة، الصوتيات، ومعالجة الإشارات.

وخلال العام 1980، وُجدت تقنية إحصائية من خلالها تم تطبيق نماذج ماركوف (HMM) لحل مشكلة التعرف على الكلام، وقد ظهرت نتائج واعدة، بشكل غير رسمي.

إن نماذج ماركوف (HMM) تقوم بمعالجة تدفق الصوت ديناميكيًا (من اليسار إلى اليمين)، باستمرار الحوسبة وتحديث احتمالية أن واحدة أو أكثر من التفسيرات هي الإجابة الصحيحة، فهذا أدى إلى العديد من التعرف على الكلام المتاح تجاريا في تسويق المنتجات، وأبرزها "الكلام الطبيعي" أو ما يعرف بـ"NaturallySpeaking" من نظم الاثنين، وقد أظهر تحسُّنًا ملحوظًا على الجهود السابقة، هذا النهج (على الأقل في سنواتها السابقة) كان لا يزال دقيقًا بشكل لا لبسَ فيه، وكان هذا اعتمادًا واسع النطاق من هذه التكنولوجيا.

وفي الآونة الأخيرة، مع تطبيق التعلم الآلي الأحدث تقنيًا، مرة أخرى يكون الروبوت مدفوع القدرة على التقاط وتحليل مجموعات كبيرة من عينات الكلام، مع زيادة الذكاء الاصطناعي، وتعيين الدقة وفائدة هذه النُظم.

وفي عام 2009، تعاون مجموعة من الباحثين في جامعة "تورونتو" مع شركة "IBM" من أجل البحث في تطبيق تقنيات التعلم الآلي للحدّ من المشكلة، ومن معدلات الخطأ، فكانت النتيجة حسنة بنسبة 30٪، وقد وُجدَ استخدامٌ رئيسيٌّ في الهواتف الذكية كطريقة بديلة لإصدار الأوامر وإدخال البيانات، وبالتالي تأججَ انفجار الاهتمام والبحث مرة أخرى، فقد أصبح الجمع بين أجهزة الكمبيوتر أكثر قوةً، وأسرع في الوصول إلى كميات

كبيرة من بيانات التدريب، وتآمرت أساليب التعلم الآلي للقضاء على المشكلة وتقديمها نظم ذات أهمية عملية وتجارية، بينما الحالة الراهنة للفن في التعرف الآلي على الكلام هي بالتأكيد أقل قدرة من البشر المتحدثين، وفائدة تقنية هذا النطاق المحدود مثيرة للإعجاب، على سبيل المثال، في Google Voice وSiri من Apple، كل منها متاح على الهواتف الذكية للشركة المعنية.

رابعًا: معالجة اللغة الطبيعية:

قدرة البشر على استخدام اللغة هي عامل مميز ورئيسي بينهم وغيرهم، فنحن نستخدم كلماتنا ليس للتواصل فقط، ولكن أيضًا لمساعدتنا على التفكير والتذكر، وتعيين الأشياء والفئات، وتسمية الأفراد.

إن اللغة لا تُستخدم فقط في الوصف، لكنها تُستخدم أيضًا للتثقيف، والإبداع، والتخيل، وتأكيد النوايا، وجعل الالتزامات، وتحديد هوية الناس تراث مماثل، من بين أشياء أخرى كثيرة، مثلنا.

إن اللغات تتطور وتتناسب مع احتياجاتنا، تقريبًا كما لو كانوا كائناتٍ حيةً في حد ذاتها، فهناك الكثير من النظريات المتنافسة حول التطور اللغوي التي ناقشتها الجمعية اللغوية في باريس في مناقشة أصول اللغة في الآونة الأخيرة، إلا أن عالم اللغويات الأسطوري "نعوم تشومسكي" (من بين آخرين) قد ناقش أصول اللغة سواء كانت اللغة تطورت على الإطلاق أو كانت نتيجة للغناء.