الذكاء الاصطناعي القابل للتفسير

جمال مراد قيس

2026-02-03

يتناول المقال أهمية **الذكاء الاصطناعي القابل للتفسير** بوصفه شرطًا أساسيًا لبناء الثقة في عصر الخوارزميات، من خلال تمكين الإنسان من فهم قرارات الأنظمة الذكية ومساءلتها، خصوصًا في المجالات الحسّاسة التي تمسّ حياة الأفراد وحقوقهم.

لم يعد السؤال اليوم: هل يعمل الذكاء الاصطناعي؟ بل أصبح السؤال الأكثر إلحاحًا: كيف ولماذا اتخذ هذا القرار؟ ففي عصر تتسلل فيه الخوارزميات إلى تفاصيل الحياة اليومية، من التشخيص الطبي والتمويل إلى التعليم والقضاء، لم تعد الكفاءة وحدها كافية لبناء الثقة. هنا يبرز مفهوم الذكاء الاصطناعي القابل للتفسير (Explainable AI – XAI) بوصفه حجر الزاوية في العلاقة المستقبلية بين الإنسان والآلة.

على مدار العقد الأخيرحققت نماذج الذكاء الاصطناعي قفزات نوعية في الدقة والأداء، خاصة مع انتشار التعلم العميق والشبكات العصبية المعقدة. غير أن هذه النجاحات جاءت بثمن معرفي خفي: تحوّل كثير من هذه النماذج إلى ما يشبه "الصندوق الأسود"، حيث تُقدَّم النتائج دون فهم واضح للمنطق الذي قاد إليها.

هذا الغموض، وإن كان مقبولًا في بعض التطبيقات الترفيهية، يصبح إشكاليًا بل وخطيرًا حين يتعلق الأمر بقرارات تمسّ صحة الإنسان، أو حقوقه، أو مستقبله المهني والاجتماعي.

الذكاء الاصطناعي القابل للتفسير لا يسعى إلى إضعاف قوة النماذج، بل إلى إعادة التوازن بين الأداء والفهم. هو محاولة علمية وأخلاقية لجعل الخوارزميات قادرة على شرح قراراتها بلغة يمكن للبشر استيعابها، سواء كانوا أطباء، أو قضاة، أو باحثين، أو حتى مستخدمين عاديين. فالتفسير هنا ليس ترفًا معرفيًا، بل شرطًا أساسيًا للمساءلة، والمراجعة، وبناء الثقة.

في المجال الطبي قد ينجح نموذج ذكاء اصطناعي في اكتشاف ورم بدقة تفوق الإنسان، لكن الطبيب لن يغامر باتخاذ قرار علاجي مصيري ما لم يفهم الأسس التي بُني عليها هذا التشخيص. التفسير يمنح الطبيب أداة للمقارنة، والتحقق، وربما الاعتراض، ويحوّل الذكاء الاصطناعي من "حكم نهائي" إلى شريك في اتخاذ القرار.

الأمر ذاته ينطبق على القطاعات المالية، حيث تؤثر الخوارزميات على منح القروض أو تقييم المخاطر، وعلى الأنظمة القانونية التي بدأت تختبر أدوات ذكاء اصطناعي لدعم الأحكام القضائية.

لكن التفسير ليس مفهومًا واحدًا بسيطًا. فهناك مستويات متعددة للفهم: تفسير تقني موجّه للخبراء، وتفسير بصري أو لغوي مبسّط للمستخدمين غير المتخصصين، وتفسير سياقي يربط القرار بالبيئة الاجتماعية والقانونية التي اتُخذ فيها. التحدي الحقيقي يكمن في تقديم التفسير المناسب للشخص المناسب، دون الإخلال بدقة النموذج أو تعقيده.

التقدم في مجال الذكاء الاصطناعي القابل للتفسير يعكس إدراكًا عالميًا متزايدًا بأن الثقة لا تُفرض بالقوة التقنية، بل تُبنى بالشفافية. ولهذا السبب، بدأت التشريعات الدولية تتعامل مع التفسير بوصفه حقًا للمستخدم، لا ميزة إضافية. فحين تؤثر الخوارزمية على حياة الإنسان، يصبح من حقه أن يعرف كيف ولماذا.

في السياق العربي، تكتسب هذه القضية بعدًا إضافيًا. فاعتماد تقنيات ذكاء اصطناعي مستوردة، دون فهم عميق لمنطقها أو انحيازاتها المحتملة، قد يؤدي إلى قرارات لا تراعي الخصوصيات الثقافية والاجتماعية. الذكاء القابل للتفسير يتيح فرصة لتوطين التقنية، وفهمها، وتعديلها بما يخدم الواقع المحلي بدل أن يُعاد تشكيل الواقع ليتلاءم معها.

كما أن التفسير يعيد الاعتبار للدور الإنساني في عصر الأتمتة. فبدل أن يُقصى الإنسان لصالح خوارزمية "أذكى"، يصبح هو الحكم الأخير، القادر على الفهم والمساءلة واتخاذ القرار الأخلاقي. بهذا المعنى، لا يحدّ الذكاء الاصطناعي القابل للتفسير من قوة الآلة، بل يحمي إنسانية القرار.

نحن اليوم على أعتاب مرحلة جديدة في تطور الذكاء الاصطناعي، مرحلة لا يُقاس فيها التقدم بعدد المعاملات الحسابية أو دقة التنبؤ فحسب، بل بقدرة الأنظمة على الشرح، والتبرير، والتحاور مع العقل البشري. وفي عالم تحكمه الخوارزميات بشكل متزايد، تصبح الشفافية ليست خيارًا، بل شرطًا أساسيًا لبقاء الثقة.

الذكاء الاصطناعي القابل للتفسير هو اعتراف ضمني بأن المستقبل لا يُبنى على التفوق التقني وحده، بل على شراكة واعية بين الإنسان والآلة، حيث يفهم كل طرف منطق الآخر، ويعملان معًا في فضاء من المسؤولية والمعرفة المشتركة.

المصادر

-https://research.ibm.com/ai/explainable IBM Research – Explainable AI

//:https DARPA – Explainable Artificial Intelligence (XAI) ai

European www.darpa.mil/program/explainable-artificial-intelligence

https://digital-strategy.ec.europa.eu Commission – Trustworthy AI

https://ai.google/research Google AI – Interpretability and Explainability

//:https MIT Technology Review – Explainable AI

www.technologyreview.com

تواصل مع الكاتب: mohamedmouradgamal@gmail.com