محاضرة 1: ما الذكاء الاصطناعي؟ وكيف "يفكّر" الحاسوب في قضايا المجتمع؟

تمهيد

يشير الذكاء الاصطناعي (AI) إلى محاولة جعل الحاسوب يؤدي مهامًا تشبه ما يفعله البشر: فهم اللغة، التعرّف إلى الأنماط، مساعدة القرار، وتوليد الصياغات النصّية. يهمّ هذا الموضوع عالم الاجتماع لأنه بات حاضرًا في التعليم والصحّة والإعلام والسياسات العامة، وبات يؤثّر—مباشرة أو عبر خوارزميات التصفية والتوصية—في الفرص والموارد والتصوّرات الاجتماعية .(Russell & Norvig, 2020; Crawford, 2021) الفكرة المركزيّة هنا أن الذكاء الاصطناعي **أداة تنبّؤ وتنظيم** وليست بديلًا عن الحكم العلمي والخبرة الميدانية؛ فهو يقترح، ونحن نقرّر ونحاسِب.(Mitchell, 2019)

(1تعريفات أساسية بلغة مبسّطة

الذكاء الاصطناعي هو المظلّة الكبرى، وضمنها يقع تعلّم الآلة (ML) الذي يسمح للنموذج بأن **يتعلّم من البيانات** بدل أن نكتب له كل القواعد يدويًا .(Domingos, 2015) بعض نماذج التعلّم الآلي تُبنى من شبكات عصبية عميقة تُسمّى "التعلّم العميق"، وهي بارعة في التقاط الأنماط المعقّدة كتمييز الصور أو فهم أجزاء من اللغة. أمّا **النماذج اللغوية الكبيرة** (LLMs)فهي نماذج نصّية تتنبّأ بالكلمة التالية وفقًا لأنماط تعلّمتها من كمّ هائل من النصوص؛ وهي جيّدة جدًّا في إعادة الصياغة والتلخيص والاقتراح، لكنها قد تُنتج مخرجات تبدو صحيحة لكنها غير مسنَدة، وهو ما يُسمّى "الهلوسة.(Bender et al., 2021; Mitchell, 2019) "

(2كيف "يفكّر" الحاسوب؟ بين القواعد والتعلّم من البيانات

تفكّر الخوارزميات بطريقتين رئيسيّتين. الأولى قائمة على **قواعد صريحة** من نوع "إذا... فـ..."، وهي مناسبة للسياسات الواضحة (مثلًا: إذا تجاوز الطالب عددًا معينًا من الغيابات تُرسل رسالة تنبيه). قوّتها في الوضوح، وحدّها في عجزها أمام التعقيد الاجتماعي وتعدّد الاستثناءات. الثانية قائمة على **التعلّم من البيانات** :نزوّد النموذج بأمثلة كثيرة مع "الإجابة الصحيحة"، فيتعلّم نمطًا عامًا يُساعده على التوقّع في أمثلة جديدة. قوّة هذا الأسلوب أنه يلتقط علاقات معقّدة قد يصعب تدوينها في قواعد ثابتة، لكن خطورته أنه قد **يفرط في التكيّف** فيحفظ أمثلة التدريب بدل أن يتعلّم قاعدة قابلة للتعميم، أو يضخّم **تحيزات** موجودة أصلًا في البيانات (Breiman, 2001; Suresh & Guttag, 2021).في البحث الاجتماعي، نختار بين النهجين بحسب طبيعة المشكلة والبيانات المتاحة ومستوى المخاطر المقبولة على الفئات المتأثرة.

(3من سؤال البحث إلى نموذج: دورة حياة مختصرة

تبدأ الرحلة بـ**صياغة سؤال واضح** وغرض اجتماعي مفهوم: ماذا نريد التنبؤ به أو شرحه؟ ومن الذي سيتأثّر بالقرار؟ .(Lazer et al., 2009) يلي ذلك تحديد **البيانات** :من جمعها؟ كيف جُمعت؟ من تمثّل ومن تُهمِّش؟ توصي الأدبيات بآليات توثيق منهجية مثل **Datasheets for Datasets**التي تشرح مصدر البيانات وحدودها وجمهورها المقصود .(Gebru et al., 2021) بعد ذلك نقسّم البيانات إلى **تدريب** و**اختبار** لفحص قدرة النموذج على **التعميم** خارج ما "شاده" في التدريب. عند التقييم، لا تكفي "الدقة" المجمّعة وحدها؛ ففي قضايا اجتماعية حسّاسة قد نهتمّ مثلًا بـ**الاستدعاء** للحالات المهمّة (ألا يُفلت المستحقّون للدعم) وبـ**الدقّة الإيجابية** (ألا نحمّل غير المستحقّين تبعات قرار خاطئ)، مع الانتباه لاعتبارات **الإنصاف** بين الفئات .(Lipton, 2018) أخيرًا، ندوّن **المسؤوليات** :ماذا يحدث إذا أخطأ النموذج؟ من يراجع؟ ما آليات التظلّم؟.(Raji et al., 2020)

(4جودة البيانات والعدالة: حيث تبدأ المشاكل وتنتهي

تحدّد جودة البيانات حدود ما يستطيع النموذج تقديمه. فإذا كانت عيّنتنا **غير ممثّلة** لفئة ريفية أو لنساء أو لمهاجرين، فإن قراراتنا الآلية قد تعيد إنتاج عدم المساواة، حتى إن بدت المؤشرات "جيّدة" رقميًا ;O'Neil, 2016) .(Crawford, 2021قد ينشأ الانحياز من **مرحلة الجمع** (من شملنا ومن استثنينا؟)، أو **المعالجة** (كيف ملأنا القيم المفقودة؟)، أو **اختيار الهدف** (هل نقيس "الجدارة" بمعايير طبقية؟)، أو من **سياق النشر** عندما نطبّق نموذجًا خارج مجاله .(Suresh & Guttag, 2021) للتخفيف، نقترح الجمع بين توثيق البيانات عبر **Datasheets**وتوثيق النماذج عبر **Model Cards**التي توضّح الجمهور المستهدف والافتراضات والقيود، إضافة إلى مراجعات أخلاقية داخلية وخارجية.(Gebru et al., 2021; Mitchell et al., 2019)

(5أين تفيدنا النماذج اللغوية الكبيرة؟ وأين تضلّل؟

تفيد النماذج اللغوية الكبيرة في دعم الباحث اجتماعيًا: اقتراح **أسئلة بحث** بديلة، بناء **مخطط مراجعة أدبيات** أولي، صياغة **مسوّدات** لدليل مقابلة أو بنود استبيان، وتنظيم **الملاحظات** وترشيح ثيمات أولية. لكن حدودها واضحة: قد **تهلوس** مراجعًا أو حقائق بلا سند؛ وقد تعكس **تحيزات لغوية وثقافية** من بيانات تدريبها؛ وقد لا تفهم السياق المحلّي أو السخرية أو الإشارات الثقافية الدقيقة .(Bender et al., 2021; Crawford, 2021) الاستخدام المسؤول يقتضي التعامل معها كـ**مساعد تحرير وتنظيم** لا كمصدر نهائي للمعلومة، مع **تحقّق بشري مستقل** قبل الاستشهاد أو اتخاذ قرار.

(6قراءة مخرجات الذكاء الاصطناعي نقديًا

يمكن تبنّي "روبرك" بسيط من أربعة أبعاد :**الملاءمة**)هل يخدم سؤالنا فعلًا؟(، **التغطية**)هل شمل المحاور الأساسية أم ترك فراغات؟(، **الدقّة الواقعية** (هل تتضمن المخرجات ادّعاءات أو مراجع تحتاج تحققًا؟)، و**الإجرائية**)هل أستطيع تحويل المخرج إلى إجراء أو أداة أو فقرة مفهومة؟ (Doshi-Velez & Kim, 2017; Lipton, 2018). إضافة إلى ذلك، ينبغي التفريق بين **الارتباط والسببية؛** فارتفاع "الرضا عن الخدمات" بالتزامن مع سياسة معيّنة لا يعني أنّ السياسة تسبّبت في الارتفاع؛ قد يوجد متغيّر وسيط أو ظرف اقتصادي أوسع (Lazer et al., 2009). القراءة النقدية ليست اعتراضًا على الذكاء الاصطناعي، بل هي أسلوب استخدامه بأمان وفعالية اجتماعية.

(7أخطاء شائعة وكيف نتفاداها

من الأخطاء المتكرّرة **الاعتماد على الدقة وحدها** في سياقات حسّاسة؛ فمثلًا في التنبؤ بالتسرب المدرسي، قد نفضّل تعظيم **استدعاء** الحالات المعرضة للخطر—حتى على حساب بعض الدقة الإجمالية—لضمان عدم تفويت مستحقّي الدعم .(Lipton, 2018) خطأ آخر هو **التعميم خارج النطاق؛** إذ يُنقل نموذج دُرّب في مدينة ساحلية غنية ليُطبَّق كما هو في منطقة ريفية بلا فحص **تحوّل التوزيع** .كما يحدث التباس عندما لا تتضح **أدوار المساءلة** :من يشرح النتائج لأصحاب المصلحة؟ من يراجع الأخلاقيات؟ ما إجراء التصحيح عند ظهور انحياز؟ ينبغي توثيق ذلك مسبقًا.(Raji et al., 2020)

(8مثال اجتماعي توضيحي

تصوّروا مشروعًا للتنبّؤ بـ**احتمال التسرّب المدرسي** .إذا جمعنا البيانات من مدارس حضرية ثريّة فقط، فسيتعلّم النموذج أنماطًا لا تمثّل الريف أو الأحياء الفقيرة. عند التطبيق على منطقة ريفية، قد **يقلّ الأداء** ويُوزّع الدعم بشكل غير عادل. الحلّ ليس رفض الذكاء الاصطناعي، بل **تحسين تمثيل البيانات،** وتوثيق الفجوات عبر **Datasheets،** وتقييم النموذج بمقاييس تُراعي العدالة، ثم نشر **Model Card**توضّح الحدود والجمهور المستهدَف.(boyd & Crawford, 2012; Gebru et al., 2021)

خاتمة

الذكاء الاصطناعي في العلوم الاجتماعية أداة قويّة لتسريع التفكير وتنظيم العمل وفتح أسئلة جديدة، لكنه يظلّ **مرآة** تعكس جودة البيانات ووضوح الأسئلة ونزاهة الإجراءات. حين نُحسن توثيق البيانات والنماذج، ونُقيم الأداء بأكثر

من رقم، ونُراعي العدالة والسياق، يصبح الذكاء الاصطناعي **مساعدًا منهجيًا** موثوقًا بدل أن يكون مصدرًا جديدًا للالتباس أو الظلم.(Mitchell, 2019; Crawford, 2021)

قائمة المراجع

1. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
2. boyd, d., & Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679.
3. Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science, 16*(3), 199–231.
4. Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
5. Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.
6. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint* arXiv:1702.08608.
7. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM, 64*(12), 86–92.
8. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., … Van Alstyne, M. (2009). Computational social science. *Science, 323*(5915), 721–723.
9. Lipton, Z. C. (2018). The mythos of model interpretability. *Queue, 16*(3), 31–57.
10. Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.
11. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229).

12. O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
13. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44).
14. Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
15. Suresh, H., & Guttag, J. V. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. In *Proceedings of the 2021 ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21)*.