



Los peligros de la certeza basada en datos

Artículo traducido del original en inglés escrito por Diane Coyle¹ y publicado en la revista Project Syndicate el Abr 14, 2021

Los economistas se apresuran a aceptar el uso del big data en sus investigaciones, mientras que muchos responsables políticos piensan que la inteligencia artificial ofrece margen para una mayor rentabilidad y mejores resultados políticos. Pero antes de confiar más decisiones a los sistemas de aprendizaje automático y IA basados en datos, debemos ser claros sobre las limitaciones de los mismos.

CAMBRIDGE - Los amigos que trabajan en las artes y las humanidades han comenzado a hacer algo inusual, al menos para ellos: analizar los datos. Esto se debe a la pandemia, por supuesto. Todos los días, comprueban el número de casos covid-19, qué tan lenta o rápidamente el factor R está disminuyendo, y cuántas personas en nuestra área se vacunaron el día anterior.

Mientras tanto, las redes sociales están llenas de reclamos y contrademandas sobre todo tipo de otros datos. ¿Está disminuyendo o aumentando la pobreza mundial? ¿Cuál es el nivel real del desempleo estadounidense? El escrutinio, que a veces conduce a argumentos tensos, es el resultado del deseo de las personas de citar -o desafiar- la autoridad de los datos para apoyar su posición o visión del mundo.

Pero en otras áreas donde se utilizan datos, hay notablemente poco enfoque en su fiabilidad o interpretación. Un ejemplo llamativo que he notado recientemente se refiere a las pruebas "CAPTCHA" diseñadas para proteger sitios web contra bots, que le piden que pruebe su humanidad mediante la identificación de imágenes que contienen características comunes como barcos, bicicletas o semáforos. Si su elección, incluso si es correcta, difiere de la del sistema de máquinas que utiliza su selección para entrenar un algoritmo de reconocimiento de imágenes, se considerará no humano.

En este ejemplo, el error de la máquina es obvio, aunque no hay ninguna apelación en su contra si desea acceder al sitio web que está guardando. Pero en otros casos, es posible que no sea posible identificar qué conclusiones están sacando los sistemas de aprendizaje automático o los analistas humanos cuando ponen más peso en los datos de lo que los datos pueden soportar.

¹ Diane Coyle, es profesora de políticas públicas en la Universidad de Cambridge, es la autora, más recientemente, de Markets, State, and People: Economics for Public Policy.



Los economistas se apresuran a aceptar el uso del big data en sus investigaciones, mientras que muchos responsables políticos piensan que la inteligencia artificial ofrece margen para una mayor rentabilidad y mejores resultados políticos. Pero antes de confiar más decisiones a los sistemas de aprendizaje automático y IA basados en datos, debemos ser claros sobre las limitaciones de los datos.

Ya se presta muy poca atención a las incertidumbres inherentes a los datos económicos. Aunque los responsables de la formulación de políticas generalmente aprecian que incluso algo tan básico como el crecimiento del PIB está sujeto a grandes incertidumbres y revisiones, parece imposible impedir que la gente construya narrativas sobre bases débiles.

Por ejemplo, las comparaciones entre países sobre el impacto de la pandemia en el PIB nacional están plagadas de dificultades, debido a las diferencias en la estructura económica y la metodología estadística. Pero eso no detiene las afirmaciones sobre qué economías están capeando la crisis mejor o peor.

O considere la "verdadera" tasa de inflación. Las disputas aparentemente técnicas sobre la mejor manera de construir un índice de precios enmascaran profundos conflictos distributales, como los entre prestatarios y tenedores de bonos, o trabajadores y empleadores.

Los datos que utilizamos dan forma a nuestra visión de un mundo complejo y cambiante. Pero los datos representan la realidad desde una perspectiva particular. Los datos del tipo desplegados en los debates políticos rara vez están completamente desenfocados del mundo que describen, pero la lente que proporcionan puede ser nítida o borrosa, y no hay escapatoria a la perspectiva que ofrecen.

Una posible razón de la desconfianza actual de la "experiencia" económica es la creciente brecha entre las evaluaciones económicas técnicas de arriba hacia abajo basadas en series de datos familiares y un mundo alternativo de datos más granulares que presentan el panorama ascendente. Las estadísticas económicas estándar capturan la experiencia promedio, que deja de ser típica cuando la fortuna de las personas diverge. ¹

En general, los defensores de la política basada en pruebas son conscientes de la incertidumbre inherente de los datos disponibles. Los investigadores se preocupan mucho con respecto al muestreo, el margen de error y las limitaciones del método de recopilación de datos utilizado. Pero el grado de falsa certeza tiende a aumentar con la proximidad a la toma de decisiones políticas y políticas.



El expresidente estadounidense Harry S. Truman está lejos de ser el único político que ha expresado impaciencia con los economistas que dicen: "'Por un lado...'", entonces, 'pero por el otro'".

Sin embargo, el hambre actual de certeza basada en datos se está volviendo peligroso a medida que nos basamos cada vez más en procedimientos de decisión tecnocráticos, incluidos los sistemas de aprendizaje automático, para la formulación de políticas en áreas como la justicia penal, la policía y el bienestar. Las democracias a menudo se basan en la ambigüedad constructiva para conciliar intereses en conflicto, como los relativos a la distribución de los retornos a un activo, o para abordar la cuestión de si las autoridades encargadas de hacer cumplir la ley deben errar del lado de encarcelar a personas inocentes o dejar que los criminales salgan libres. Las reclamaciones a la autoridad basada en datos minimizan o eliminan el alcance de la ambigüedad, con consecuencias potencialmente significativas.

Estoy a favor de más y mejores datos, que han sido esenciales para los esfuerzos de los gobiernos para gestionar la pandemia. Pero cuanto más usamos los datos para tomar decisiones, más sensibles debemos ser al hecho de que los datos pintan la vista de un experto o máquina, basada en categorías ideadas por alguien que es a su vez un jugador en el juego de estatus de la sociedad. De lo contrario, terminaremos con procesos de decisión al igual que esas pruebas CAPTCHA pícaras, insistiendo en que un barco es una bicicleta, y dejando a otras personas sin más opción que ponerse de acuerdo.