

Prof. Dr. Andrea Castaldo

Catedrático de Derecho Penal. Univ. de Salerno. Abogado, Salerno, Italia. Socio de la FICP.

~The noise (el rumor): el desafío de la A.I. (inteligencia artificial) entre predicción y manipulación de la realidad~

La previsibilidad del resultado de una disputa -o de su posible ocurrencia- responde ciertamente al principio de expectativas legítimas. En el contexto que hoy nos ocupa, cabe preguntarnos si es posible poder predecir una sentencia de un juez con una buena dosis de certeza.

Bien puede suceder que dos jueces tomen decisiones diferentes para casos similares. Esto sucede por el "ruido" o por las variables en los juicios que abstractamente deberían ser idénticos, lo cual resulta en una pérdida de verdad en el juicio. "Nihil novi sub sole": todos cometen errores en la evaluación.

La psicología del comportamiento ya ha estudiado este fenómeno y ha sido definida en su valor jurídico y penal con el llamado "sesgo cognitivo" u oblicuidad, las distorsiones que cada uno de nosotros, conscientemente o no, aplicamos en el juicio.

Los "sesgos cognitivos" son construcciones basadas, fuera del juicio crítico, en percepciones erróneas o deformadas, en prejuicios e ideologías. La mayoría de las veces, estos son errores cognitivos impactan en la vida cotidiana, no solo en las decisiones y comportamientos, sino también en los procesos de pensamiento.

A esto, hoy, se suma el "ruido", que sistemática e inevitablemente produce una distorsión.

El "ruido" identifica la variación de juicios que deberían ser idénticos entre sí. Se diferencia del "sesgo" que representa más genéricamente el prejuicio, el error direccional predecible en las decisiones humanas.

El ruido es una variabilidad no deseada en el juicio profesional. El sesgo, por otro lado, es más bien la tendencia de un individuo a utilizar los mismos patrones de toma de decisiones en situaciones similares. A pesar de las grandes diferencias entre estos dos conceptos, ambos representan errores de juicio.

Para comprender mejor cómo y cuánto afecta realmente el "ruido", puede ser útil el estudio realizado por el premio Nobel de economía Daniel KAHNEMAN, quien demostró la existencia del "ruido" en todo juicio humano.

El estudio ha evidenciado cómo nuestros juicios, incluso los profesionales, están condicionados por este "ruido". En las salas de audiencias, por ejemplo, se espera que el juez exprese su decisión

con objetividad, siguiendo criterios objetivos de evaluación. Kahneman, por su parte, mostró que, incluso en sede judicial existe el "ruido" que contamina el trabajo del juez y sus convicciones.

Para ello, citó el siguiente experimento: 50 jueces estadounidenses de diferentes distritos se involucraron y se les pidió que decidieran los mismos casos. Sorprendentemente, por el mismo hecho, un juez dictó una pena de veinte años de prisión, mientras que otro juez optó por una pena de sólo tres años de prisión.

Para comprender si el resultado diferente se debe a la incidencia del "ruido", un estudio ha demostrado cómo el estado de ánimo de los jueces puede influir en sus juicios. De hecho, se ha señalado que, en los días posteriores a la derrota del equipo de la Ciudad, las sentencias de los jueces son más severas. Por el contrario, en los días posteriores a una victoria sus sentencias son más indulgentes.

En este contexto, podría ser legítimo, o al menos natural, pensar que la Inteligencia Artificial, los algoritmos, impermeables a las distorsiones humanas, son más fiables.

Hoy en día, los algoritmos predictivos se utilizan en los más diversos campos para garantizar un juicio más justo y menos variable.

Hay algoritmos que pueden predecir las habilidades y confiabilidad de los candidatos a un puesto de trabajo, simplemente con base en entrevistas de video, otros que pronto podrían usarse en las escuelas para predecir el nivel de atención y educación de los estudiantes.

Los complejos sistemas de inteligencia artificial ya se utilizan en varias partes del mundo para autos sin conductor, diagnósticos médicos, perros policías robóticos e incluso armas "inteligentes".

Todo esto puede dar lugar a escenarios distópicos, como el sistema de "puntuación social" -es decir, la asignación de una puntuación y premios relacionados en función de un comportamiento "virtuoso"- adoptado por China, en el que un algoritmo puede clasificar a los ciudadanos en función de cuánto respeta las reglas y cuánto muestra obediencia (en línea y fuera de línea) al partido. Otros riesgos, mucho más concretos, son la discriminación en el ámbito laboral por razón de sexo u orientación sexual o en perjuicio de determinadas minorías sociales en contextos de policía predictiva.

Por otra parte, el control de las emociones por parte de los sistemas de IA puede conducir a la manipulación del comportamiento de niños, ancianos u otros consumidores, explotando sus vulnerabilidades cognitivas e induciéndolos a elecciones comerciales no deseadas. Al mismo tiempo, estos sistemas automatizados suelen ser opacos y, por lo tanto, difíciles de impugnar.

Ahora bien, queda sin respuesta la pregunta de si el juicio humano o el juicio basado en sistemas de Inteligencia Artificial es más confiable.

**Actas del III Congreso Internacional de la FICP, Alcalá de Henares, Madrid (España),
septiembre de 2022**

Si, como se mencionó, el "ruido" afecta nuestros juicios y nuestras elecciones, cabe preguntarnos si resulta posible que también afecte las estrategias en la base de la lucha contra los fenómenos criminales. Tomemos por ejemplo la lucha contra la corrupción.

En la medida en que afecta la percepción del fenómeno de la corrupción, el "ruido" puede conducir a la sobreproducción de normas destinadas a prevenir y combatir la corrupción.

Sin embargo, de esta manera terminamos amplificando la idea de un fenómeno en constante expansión que no puede ser derrotado, también gracias al énfasis de los medios en la difusión de los casos de corrupción.

En estos términos, como podéis imaginar, una percepción distorsionada del fenómeno afecta también a las políticas de prevención, que siguen repitiéndose para intentar dar respuesta a la "petición de hacer algo".

La aparente ineficacia de las medidas preventivas condicionadas, como se mencionó, por el "ruido", termina por desviar las esperanzas de constatar el fenómeno en los jueces, quienes pueden sentirse con derecho a imponer severas sanciones para combatir el fenómeno. En definitiva: el "ruido" puede afectar al sistema anticorrupción tanto en la fase de análisis y percepción del fenómeno, como en las estrategias de prevención y represión.

Llegados a este punto, intentemos comprobar si la Inteligencia Artificial puede ayudar a limitar o eliminar el efecto del "ruido" en la lucha contra la corrupción.

Como es bien sabido, la Inteligencia Artificial ha sido definida por Oxford Insights como "la próxima frontera de la anticorrupción".

En primer lugar, la contribución de la máquina puede ayudar a medir la corrupción de una manera más objetiva, desvinculándola de factores subjetivos como la percepción.

La identificación de indicadores comunes y objetivos para medir la corrupción es el primer paso fundamental para mejorar las estrategias de prevención y represión.

En segundo lugar, la inteligencia artificial se puede utilizar para identificar los síntomas de corrupción antes de que ocurra. El estudio de la Escuela Superior de Económicas y la Universidad de Valladolid consiguió aislar los siguientes factores que conducen a la corrupción pública: la tributación de la propiedad, las subidas de precios y la presencia del mismo partido en el poder.

La investigación ha comprobado que la Inteligencia Artificial, gracias al análisis de estos datos, puede predecir el fenómeno de la corrupción hasta tres años antes de que se produzca. Desde el punto

**Actas del III Congreso Internacional de la FICP, Alcalá de Henares, Madrid (España),
septiembre de 2022**

de vista de la represión, según Kahneman, la combinación de algoritmos con la decisión del juez sobre el monto de la pena puede reducir o eliminar las distorsiones por "ruido".

El uso del algoritmo en la decisión judicial puede generar sentencias más uniformes, independientemente del contexto emocional y "ruido" del juez individual. Menos discrecionalidad en la decisión garantizaría juicios más objetivos y menos contaminados por "ruido".

Una mayor objetividad en la aplicación de la sentencia ayudaría no solo al juez a tomar decisiones sin "ruido", sino que sería una forma de control para los mismos ciudadanos que, con 'algoritmos judiciales', tendrían mayor certeza en las condenas penales.

En conclusión, existen numerosos beneficios potenciales derivados de la aplicación de la Inteligencia Artificial a las estrategias anticorrupción.

Sin embargo, dado que la máquina aún no es capaz de aprender y generar resultados de forma completamente autónoma, habrá que evitar que los insumos que se le den a la máquina que ha de generar la decisión sean estropeados por el "ruido" humano.

Con otras palabras, el riesgo de confiar excesivamente en el potencial de la máquina podría subestimar el peligro de que la información que le transfieren los humanos se contamine con valoraciones subjetivas y parciales.

Por tanto, sólo la adecuada formación y "entrenamiento" de quienes tendrán la tarea de manejar la máquina o su total automatización puede garantizar el inicio de una nueva era anticorrupción.