Los algoritmos no experimentan alegría, miedo, tristeza o indignación. Sin embargo, pueden inferir qué contenidos nos alteran, predecir cómo reaccionaremos y reorganizar nuestro entorno digital para mantenernos conectados. Comprender este circuito es fundamental para proteger nuestra autonomía emocional.

[ILUSTRACIÓN: ALENA IVOCHKINA / ISTOCK]

A las once y media de la noche, una persona abre una red social después de un día difícil. Se detiene ante un vídeo sobre agotamiento, vuelve a ver otro sobre rupturas y termina leyendo una historia triste. A la mañana siguiente, cuando ve que su pantalla parece haberse llenado de ansiedad, nostalgia y conflictos, piensa: "El algoritmo sabe cómo me siento".

En realidad, no lo sabe como una persona cercana. No comprende el cansancio, la tristeza o el miedo. Hace algo distinto: observa conductas, calcula probabilidades y aprende qué estímulos consiguen detenernos. No necesita acceder a nuestra vida interior para intervenir en ella. Le basta con registrar qué miramos, cuánto tiempo permanecemos, qué repetimos, qué compartimos y cuándo regresamos.

El algoritmo no lee la mente: lee el comportamiento

Cada interacción deja una señal: pulsar "Me gusta" es la más evidente, pero no necesariamente la más informativa. También cuentan el tiempo de visualización, la velocidad de desplazamiento, las búsquedas, los comentarios y los contenidos que abandonamos.

Con esos rastros, los sistemas construyen predicciones. Un trabajo sobre señales afectivas en recomendadores de redes sociales mostró que es posible anticipar si una publicación resultará entretenida, inspiradora, informativa o irritante. Otros estudios han comprobado que el lenguaje compartido en Facebook puede aportar información para predecir posteriormente indicadores de depresión registrados en historias clínicas.

Esto no significa que la tecnología diagnostique de manera infalible nuestras emociones. La emoción depende del contexto, la cultura y la historia personal. Una sonrisa puede expresar alegría, cortesía, nerviosismo o disimulo. Una revisión exhaustiva sobre reconocimiento facial advirtió que los movimientos del rostro no permiten identificar estados emocionales universales con la precisión que a menudo promete la industria. Los algoritmos trabajan con probabilidades: clasifican señales, pero no experimentan sentimientos.

Los algoritmos no experimentan emociones: calculan qué estímulos aumentarán la probabilidad de una reacción

El paso decisivo se produce cuando la predicción se convierte en recomendación. Las plataformas no muestran todo lo publicado: ordenan miles de contenidos y deciden cuáles van a ocupar los primeros segundos de nuestra atención. Esa selección modifica el entorno emocional que habitamos.

Lo negativo dispone de una ventaja competitiva. Un análisis de más de 370 millones de impresiones de titulares encontró que las palabras negativas aumentaban la probabilidad de hacer clic. El resultado no demuestra que queramos vivir enfadados o asustados, sino que la amenaza, el conflicto y la pérdida captan con rapidez nuestra atención.

La exposición tampoco es emocionalmente neutra. Un experimento con usuarios de Facebook, mostró cómo, al reducirse la presencia de publicaciones positivas o negativas en el muro de noticias de un usuario, se producen luego pequeños cambios en el lenguaje emocional de lo que este usuario publica después.

En su momento, este estudio generó un debate ético al alterar los contenidos del muro de noticias para provocar una reacción emocional, pero también mostró que la arquitectura de una plataforma puede influir en la expresión afectiva de sus usuarios.

La indignación viaja más deprisa: lo que nos hace interactuar no siempre coincide con lo que preferimos ni con lo que nos hace bien

No todas las emociones tienen el mismo rendimiento algorítmico. Las que generan alta activación –indignación, miedo, entusiasmo o asombro– favorecen la respuesta inmediata. En redes políticas, el lenguaje moral y emocional se difunde más que el neutral. Otro estudio encontró que la hostilidad hacia el grupo contrario era un predictor especialmente potente de interacciones.

El problema aparece cuando el sistema aprende que la confrontación funciona. Una auditoría publicada en 2025 comparó un ranking basado en *engagement* con un orden cronológico. El algoritmo amplificaba contenido emocionalmente cargado y hostil hacia el adversario político. Lo más revelador es que los usuarios no afirmaban preferir ese contenido cuando se les preguntaba directamente.

Lo que nos hace reaccionar no siempre coincide con lo que valoramos. Podemos detenernos ante una humillación, discutir o compartir una noticia irritante y considerar que una experiencia nos ha hecho sentir peor. Sin embargo, para un sistema que interpreta cada reacción como interés, la protesta puede parecer entusiasmo y el enfado, satisfacción.

Una tristeza pasajera puede convertirse en una dieta informativa repetitiva

La relación entre persona y algoritmo no avanza en una sola dirección. Elegimos contenidos según nuestro estado de ánimo, pero esos contenidos también pueden modificarlo. Un estudio formado por cuatro experimentos encontró una relación bidireccional: las personas con peor estado emocional consultaban información más negativa y esa navegación, a su vez, empeoraba el ánimo.

El recomendador puede intensificar el circuito. Si una persona triste se detiene ante contenidos tristes, el sistema aprende que esa temática "funciona" y ofrece más. La plataforma devuelve una versión concentrada de un estado quizá transitorio. Lo que empezó como una noche mala puede convertirse en una dieta informativa repetitiva.

La investigación sobre interacción entre humanos e inteligencia artificial ha mostrado además que pequeños sesgos pueden ser adoptados y amplificados por los usuarios. En una serie de experimentos con 1 401 participantes, las personas modificaron sus juicios emocionales y sociales después de interactuar repetidamente con modelos sesgados, a menudo sin ser conscientes de esa influencia.

El problema es qué objetivo se le ha pedido maximizar al algoritmo: si el éxito se mide en clics, el sistema busca contenidos que los impulsen

Hablar de manipulación algorítmica puede llevarnos a imaginar una máquina que desea enfadarnos. Esa imagen es incorrecta. El algoritmo no tiene intenciones: optimiza una función. El verdadero problema es qué se le ha pedido maximizar.

Cuando el éxito se mide mediante clics, minutos de uso, comentarios o retornos, el sistema busca contenidos

capaces de producir esas conductas. La emoción se convierte en un recurso económico: no porque la plataforma conozca nuestra verdad psicológica, sino porque necesita anticipar nuestra próxima acción.

Esto explica por qué la personalización puede sentirse íntima aunque sea estadística. El sistema no sabe quiénes somos; sabe qué estímulos aumentan la probabilidad de que continuemos conectados. Confundir *engagement* con bienestar equivale a confundir que alguien no pueda apartar la mirada con que esté disfrutando.

La autonomía digital también exige decidir qué entorno emocional queremos habitar

La misma capacidad puede utilizarse con otros fines. Los recomendadores podrían priorizar contenidos que los usuarios consideren valiosos, introducir diversidad emocional o permitir elegir entre objetivos: informarse, relajarse, aprender o distraerse. El estudio sobre señales afectivas en sistemas de recomendación citado anteriormente logró reducir contenidos que infringían normas y aumentar aquellos que los usuarios consideraban útiles.

En salud mental existen resultados prometedores, aunque limitados. Un ensayo aleatorizado con jóvenes mostró que un agente conversacional basado en principios de terapia cognitivo-conductual podía reducir síntomas depresivos tras dos semanas de interacción. Estas herramientas no sustituyen a un profesional, pero muestran que la personalización emocional no tiene por qué orientarse exclusivamente a retener la atención.

La alternativa no es eliminar los algoritmos, sino discutir sus objetivos. Necesitamos saber por qué vemos un contenido, disponer de controles comprensibles, diferenciar preferencias declaradas de reacciones impulsivas y evaluar los efectos sobre el bienestar, no solo sobre el consumo.

Durante años hemos hablado de privacidad como el derecho a proteger nuestros datos. En la era de la recomendación algorítmica, debemos ampliar esa idea. También importa qué se deduce de esos datos, cómo se utilizan nuestras vulnerabilidades y qué ambientes emocionales se construyen a nuestro alrededor.

La alfabetización digital ya no consiste únicamente en detectar noticias falsas o proteger una contraseña. También exige preguntarnos por qué una plataforma insiste en mostrarnos aquello que nos indigna, entristece o reconforta. Y exige que las empresas asuman que una emoción no es simplemente una señal de rendimiento.

Los algoritmos quizá no sepan cómo nos sentimos. Pero pueden aprender qué nos altera, ofrecérnoslo de nuevo y convertir nuestra reacción en el dato que alimentará la siguiente recomendación. Recuperar el control comienza por comprender ese circuito y reclamar que la tecnología no solo compita por nuestra atención, sino también por merecerla.