



CARRERA: TECNICATURA SUPERIOR EN CIENCIA
DE DATOS E INTELIGENCIA ARTIFICIAL

MATERIA: TALLER DE COMUNICACIÓN

AÑO: 2025

FECHA EXAMEN: 18/08

FECHA ENTREGA: 10/10

MODALIDAD: ENSAYO

Nº ENTREGA: 1

LOS “PELIGROS” DE LA INTELIGENCIA ARTIFICIAL

“Hoy, nuestras máquinas siguen siendo creaciones sencillas, que requieren el cuidado parental y la atención vigilante de todo recién nacido, difícilmente merecedoras de la palabra ‘inteligente’. Pero dentro del próximo siglo madurarán hasta convertirse en entidades tan complejas como nosotros, y eventualmente en algo que trascienda todo lo que conocemos —en quienes podremos sentirnos orgullosos cuando se refieran a sí mismos como nuestros descendientes.”
— Hans Moravec, *Mind Children* (1988).

La inteligencia artificial ya no pertenece al futuro: habita nuestro presente más íntimo. Está en el celular que corrige lo que escribimos, en el algoritmo que decide qué noticias vemos y en la voz sintética que nos da direcciones o nos responde preguntas. Muchos la perciben como una promesa casi mágica, capaz de resolverlo todo; otros, como una amenaza que podría reemplazar al pensamiento humano. Lo cierto es que, en menos de una década, la IA dejó de ser una curiosidad científica para convertirse en una infraestructura invisible de la vida cotidiana. Algunos lo utilizan para programar, para corregir textos, o incluso como un consultor personal. Pero ¿entendemos realmente cómo funciona, o simplemente la usamos con una fe ciega que roza la dependencia?

Definiciones básicas

La base de la IA son las redes neuronales y el procesamiento de lenguaje. El **perceptrón** fue uno de los primeros sistemas en funcionar como la neurona, la unidad funcional de nuestro cerebro. Este sistema es capaz de clasificar información en categorías.

El **Procesamiento del Lenguaje Natural**, o NLP por sus siglas en inglés, es una rama de la inteligencia artificial que se enfoca en la interacción entre las computadoras y el lenguaje humano. A lo largo de las últimas décadas, esta disciplina ha evolucionado desde simples modelos estadísticos hasta redes neuronales profundas que hoy impulsan asistentes inteligentes y traductores automáticos.

Comenzamos con lo más básico: las **expresiones regulares** y la **tokenización**. Las expresiones regulares permiten buscar patrones en textos, y la tokenización es el proceso de dividir un texto en unidades significativas, como palabras o frases. Estas tareas son esenciales para que un sistema entienda cómo está estructurado un lenguaje. Se usan herramientas simples de Unix o técnicas más avanzadas como la tokenización subpalabra, lematización y segmentación en oraciones. Todo esto se complementa con la métrica de *distancia de edición*, que mide qué tan diferente es una palabra de otra, algo clave en tareas como corrección ortográfica.

Luego pasamos a los modelos de lenguaje **N-gram**. Estos modelos predicen la probabilidad de una palabra basándose en las palabras anteriores. Se evalúan con métricas como la *perplejidad* y pueden generar texto mediante muestreo de secuencias. Sin embargo,

enfrentan el problema del sobreajuste y la escasez de datos, que se corrige con técnicas como el suavizado y la interpolación.

En la clasificación de texto, el modelo de **Naive Bayes** ha sido fundamental. Es simple pero efectivo, y se usa en tareas como análisis de sentimientos y detección de spam. Este modelo estima la probabilidad de una categoría dada una palabra, y se evalúa usando medidas como precisión, recall y F1-score. También se consideran aspectos éticos y posibles sesgos al aplicar estos modelos.

La **regresión logística** es otro enfoque clásico. Utiliza la función sigmoide para modelar probabilidades y se entrena con métodos como descenso de gradiente y regularización. La pérdida por entropía cruzada y su gradiente son claves para optimizar el modelo. La regresión logística se extiende al caso multiclase y permite interpretar con claridad las relaciones entre entradas y salidas.

El siguiente paso es la **semántica vectorial**. Aquí representamos palabras como vectores en un espacio matemático. Usamos técnicas como TF-IDF, PMI y, más recientemente, *Word2Vec*. Estas representaciones permiten comparar similitudes semánticas entre palabras con métricas como el coseno. Además, revelan relaciones semánticas y pueden usarse para visualizar y analizar sesgos en el lenguaje.

Las redes neuronales traen una revolución al NLP. Los modelos *feedforward* pueden clasificar texto, pero pronto se vio la necesidad de manejar la secuencia de palabras, lo cual llevó al uso de *redes neuronales recurrentes* (RNN). Las RNN permiten recordar información previa, crucial en el lenguaje. Sin embargo, tenían problemas con secuencias largas, lo cual fue resuelto en parte con las redes LSTM, que pueden aprender dependencias a largo plazo.

El gran salto se dio con el modelo **Transformer**. Su mecanismo de atención permite que cada palabra mire a todas las otras del texto simultáneamente, lo que mejora drásticamente la eficiencia y la calidad. Este modelo eliminó la necesidad de procesamiento secuencial, permitiendo mayor paralelización. Las entradas se codifican con *embeddings* de posición y token, y el modelo aprende a generar texto con alta coherencia contextual.

Los **Modelos de Lenguaje de Gran Escala** (LLM), como GPT o BERT, usan Transformers como base. Estos modelos se entrenan con enormes cantidades de texto, en un proceso conocido como *preentrenamiento*, y luego se ajustan a tareas específicas mediante *fine-tuning*. Se evalúan con nuevas métricas y técnicas de muestreo, y su entrenamiento requiere recursos masivos. Aunque poderosos, presentan riesgos como la generación de desinformación o la reproducción de sesgos.

Los modelos enmascarados como BERT ofrecen otra forma de entrenamiento. En lugar de predecir la siguiente palabra, predicen palabras ocultas dentro de un contexto. Esto permite aprender representaciones más ricas y útiles para tareas como clasificación, reconocimiento de entidades nombradas y análisis gramatical. También introducen *embeddings contextuales*, que adaptan el significado de una palabra al contexto en el que aparece.

Por último, entramos en el terreno de la **alineación de modelos** y el **aprendizaje en contexto**. Técnicas como el *prompting* permiten controlar el comportamiento del modelo

mediante instrucciones. Métodos como *Chain-of-Thought* y *Reinforcement Learning from Human Feedback* buscan alinear los modelos con valores humanos y mejorar la calidad de las respuestas. Se investiga también cómo optimizar automáticamente los prompts y evaluar estas respuestas de manera coherente.

¿Qué es la Inteligencia Artificial?

La respuesta más sencilla es que es un sistema informático, cuyo objetivo es simular la inteligencia. Hasta ahora la considerábamos propiedad de los animales, principalmente de los mamíferos, siendo los humanos su “máximo exponente” (aunque es sabido que incluso las plantas muestran algún tipo de inteligencia).

Una definición ampliamente aceptada es “la capacidad de resolver problemas”. La inteligencia así, se vuelve **las** inteligencias, dado que cada problema tiene un área del conocimiento humano en el cual se presenta (lógico-matemática, musical, espacial).

En los primeros capítulos de *Speech and Language Processing* (Jurafsky & Martin, 2025), los autores explican que los sistemas de IA actuales (como los que generan texto o imágenes) no “comprenden” el lenguaje del modo en que lo hacen las personas: solo aprenden patrones estadísticos a partir de grandes volúmenes de datos. Es decir, reconocen las formas externas del pensamiento humano, pero no su significado profundo. Esta diferencia, aparentemente técnica, tiene un trasfondo filosófico crucial: las máquinas pueden imitar la inteligencia, pero no la experiencia. Podemos enseñarles a hablar, pero no a comprender lo que dicen.

Acá entra en juego la observación del investigador Hans Moravec (hoy conocida como *la paradoja de Moravec*): las tareas que para la mayoría de los humanos son difíciles (como el razonamiento lógico o el cálculo matemático) resultan simples para una computadora; mientras que las que nos parecen naturales (como reconocer una cara, entender una emoción o incluso la motricidad fina) son extremadamente complejas para una máquina. Lo que nos hace humanos, entonces, no es la capacidad de procesar datos, sino la de atribuirles sentido.

Inteligencia vs. Consciencia

En “El problema difícil de la Consciencia”, Chalmers divide los “problemas” de la consciencia en dos: los fáciles y el difícil. Los fáciles son aquellos que ya fueron superados por la filosofía y las ciencias: por ejemplo, distintas actividades físicas activan distintas áreas cerebrales. Pero el problema difícil de la consciencia es cómo la actividad cerebral produce la experiencia subjetiva. Concluye que a día de hoy, la consciencia no fue definida de manera completa porque nadie pudo definir ni obtener pruebas empíricas de lo que se conoce como *qualias* (el contenido de las experiencias subjetivas), en la terminología de Dennet. Desde mi punto de vista, esta ausencia de definición nos impide determinar si un sistema artificial (incluso si este muestra inteligencia por ejemplo superando el “test de Turing”) tiene consciencia o no. Entonces, cómo podemos fiarnos de algo que creemos que

es como nosotros, pero en esencia es algo distinto? Aunque en un futuro cercano adquieran cuerpo físico y se masifique la producción de robots que piensan, ¿pueden estos llegar al nivel de un humano en tanto a actividades como el pensamiento y la creatividad?

El peligro no reside tanto en la inteligencia artificial si no en la *artificialización de nuestra inteligencia*: en delegar el juicio, el criterio y la curiosidad en sistemas que operan como cajas negras. Cuando dejamos que una IA nos diga qué leer, qué mirar o qué opinar, comenzamos a perder el hábito de pensar por nosotros mismos. El filósofo Michel Foucault advertía que las sociedades modernas ejercen el poder no tanto por la fuerza, sino por la vigilancia constante. En ese contexto, los algoritmos se convierten en una forma contemporánea del *panóptico*: vemos solo lo que el sistema nos permite ver, mientras él observa todo.

George Orwell había anticipado algo similar en 1984. Cerca del comienzo de la novela, Winston Smith toma una pluma y, luego de mucho tiempo de obediencia silenciosa, decide

“Lo que ahora se disponía Winston a hacer era abrir su Diario. Esto no se consideraba ilegal (en realidad, nada era ilegal, ya que no existían leyes), pero si lo detenían podía estar seguro de que lo condenarían a muerte, o por lo menos a veinticinco años de trabajos forzados. Winston puso un plumín en el portaplumas y lo chupó primero para quitarle la grasa. La pluma era ya un instrumento arcaico. Se usaba rarísimas veces, ni siquiera para firmar, pero él se había procurado una, furtivamente y con mucha dificultad, simplemente porque tenía la sensación de que el bello papel cremoso merecía una pluma de verdad en vez de ser rascado con un lápiz tinta. Pero lo malo era que no estaba acostumbrado a escribir a mano. Aparte de las notas muy breves, lo corriente era dictárselo todo al hablescribe, totalmente inadecuado para las circunstancias actuales. Mojó la pluma en la tinta y luego dudó unos instantes. En los intestinos se le había producido un ruido que podía delatarle. El acto trascendental, decisivo, era marcar el papel. En una letra pequeña e inhábil escribió: ‘4 de abril de 1984’”

Ese gesto —escribir a mano, fuera de la mirada del Estado— se convierte en su primer acto de libertad. El solo hecho de escribir a mano activa en su cerebro conexiones que escapan al adiestramiento recibido durante años: por un instante, Winston vuelve a pensar sin la mediación del sistema. No es casual que a partir de ese momento, el protagonista tenga sueños y pensamientos recurrentes que hasta entonces no había experimentado. En esa simple acción, mecánica y humana, resurge algo que las máquinas no pueden replicar: la experiencia interior de la conciencia. Hoy, en la era digital, el equivalente sería desconectarnos por un instante de los sistemas que todo lo registran y permitirnos pensar sin mediaciones automáticas. No se trata de rechazar la tecnología, sino de recuperar la distancia crítica necesaria para usarla sin convertirnos en su extensión.

La inteligencia artificial, usada con prudencia, puede ser una herramienta poderosa para expandir nuestras capacidades. Pero si la adoptamos sin comprenderla, corremos el riesgo de que, poco a poco, se vuelva ella quien piense —o al menos, quien decida— por nosotros. Y ese sería, quizás, el verdadero peligro: no que las máquinas se vuelvan humanas, sino que los humanos olviden cómo pensar sin ellas.

Otro peligro bien conocido es la capacidad de generar contenido aparentemente real: fotos, audios y artículos periodísticos pueden ser creados sin fundamento. Esto nos conduce a

mantener una alerta nunca antes vista: todos vamos a tener que adaptarnos para ser capaces de discernir qué es IA y qué no.

Conclusiones

La inteligencia artificial no es un concepto distante ni abstracto: es parte de nuestra vida cotidiana y, al mismo tiempo, un espejo de nuestra propia inteligencia. A través de sus modelos y algoritmos, la IA imita patrones humanos sin experimentar significado ni consciencia, lo que nos recuerda que la inteligencia no es sinónimo de experiencia. Los avances tecnológicos, desde redes neuronales hasta modelos de lenguaje de gran escala, nos muestran tanto el potencial de estas herramientas como los riesgos inherentes a su uso indiscriminado.

El verdadero peligro no reside en que las máquinas se vuelvan inteligentes, sino en que los humanos deleguemos en ellas nuestra capacidad de pensar, discernir y crear. La dependencia de sistemas que operan como cajas negras amenaza nuestra autonomía cognitiva y nos expone a sesgos invisibles y manipulación. Recuperar el pensamiento crítico, mantener la distancia ética y comprender cómo funcionan estas tecnologías es imprescindible para que la IA sea un instrumento de ampliación de nuestras capacidades, y no un sustituto de nuestro juicio.

En definitiva, la inteligencia artificial puede ser una aliada poderosa si aprendemos a usarla con conocimiento y prudencia. La libertad y la consciencia humanas dependen, más que de la sofisticación de las máquinas, de nuestra decisión de mantener siempre activa la capacidad de pensar por nosotros mismos.