

Visión Centennial de la Inteligencia Artificial

Revisión y coordinación: Sergio Viera <sviera@frd.utn.edu.ar>, Celeste Torczyński <mtorczyński@frd.utn.edu.ar>

Resumen

Este es un trabajo realizado en conjunto por el curso de Análisis de Sistemas de Información de la UTN-FRD para analizar y exponer las conclusiones de la carta recientemente publicada en <https://futureoflife.org/cause-area/artificial-intelligence/> [1] donde se expresan las dudas y riesgos que el desarrollo de la Inteligencia Artificial puede traer.

Autores

Jerónimo Aita <jeroaitagm@gmail.com>, Facundo Nicolas Almada <facu9575@gmail.com>, Ignacio Javier Ancherama <ancherama.ignacio@gmail.com>, Nicolas Emilio Angerossa <nicolasemilioangerossa@gmail.com>, Aitana Arostegui <aituarostegui@gmail.com>, Juan Martín Avila <avilajuanma75@gmail.com>, Mariano Walter Baldovino <mariano1baldovino@gmail.com>, Juan Bautista Ballini <jballini@hotmail.com.ar>, Matías Bonesi <matiasbonesi@gmail.com>, Bruno Nahuel Camino <bcamino22@gmail.com>, Manuel Campos <manucampos883@gmail.com>, Nahuel Carrizo <nahuelcarrizo711@gmail.com>, Jerónimo Casañas Onganía <jeronimo.casanas@gmail.com>, Agustín Ezequiel Casco <Agustincasco99@hotmail.com>, Gerónimo Nicolás Centuperis <geronimonico31@gmail.com>, Ramiro Chicolonea <ramichoco22@gmail.com>, Gonzalo Agustín Contrera <lguss.ac@gmail.com>, Lucas Delgado <mobile.lucas18@gmail.com>, Priscila Belen Della Vecchia <prisciladellavecchia@gmail.com>, Tiago Di Stefano <tiagodistefano07.03@gmail.com>, Francisco Miguel Fajardo Figueroa <franciscofajardo190603@gmail.com>, Pablo Gabriel Fajardo Figueroa <shaguicontact@gmail.com>, Santiago Ferrari <santiagoferrari.02@gmail.com>, Elias Gabriel Garcia <eliagarcia45.elg@gmail.com>, Agustín Gomez <agusgomez82010@gmail.com>, Juan Ignacio Iacaruso <juaniaca@hotmail.com>, Maximiliano Nicolas Jarak <maxijarak+utn@gmail.com>, Joaquin Landra <joacolan986@gmail.com>, Bianca Ornella Ledesma <orneledesma15@gmail.com>, Fiorella Valentina Licet Gomez <flicet@frd.utn.edu.ar>, Samuel Maximiliano Martínez López <smartinezlopez117@gmail.com>, Alejo Montesino <alejmontesino4@gmail.com>, Martín Leonardo Páez <martin.leo.14@hotmail.com>, Axel David Palavecino <axel2002palavecino@gmail.com>, Nicolas Alejandro Perez <nicosektor@gmail.com>, Tomás Liam Pertile <tomaspertile340@gmail.com>, Tomas Pozo

Isaurralde <tomaspozoisaurralde@gmail.com>, Sofia Nicole Ramos <ramos.sofia@hotmail.com>, Rebecca De Jesus Rodriguez Enrique <beckyroen@gmail.com>, Agustín Sanchez <agustinsanchezp123@gmail.com>, Alvaro Valentin Signorio <signorioalvaro@gmail.com>, Sebastián Nicola Stefanelli Zambrano <sebas.stefanelli@gmail.com>, Lautaro Stüve <lautarostuve@gmail.com>, Joaquin Tomas Terzano <joaquin.l.terzano@gmail.com>, Diego Emanuel Torres <diegotrs93@gmail.com>, Milagros Belén Trotta <milagrostrotta1608@gmail.com>, Francisco Veiga <francisco2.veiga@gmail.com>, Sergio Gabriel Viera <sgviera@frd.utn.edu.ar>.

Introducción

Desde el comienzo de la historia la toma de decisiones estuvo basada en una cualidad: La intuición. Todas las decisiones “correctas” ocurrían de aquellos que evaluaban los riesgos y prevenían cualquier tipo de peligro para su propia supervivencia, el que tomaba las mejores decisiones siempre fue el que más experiencia poseía. Al menos así fue hasta que, como sociedad, comenzamos a tomar en cuenta la información como núcleo de nuestras elecciones. Desplazamos la experiencia personal por una recopilación de información general que nos orientaba y nos permitía minimizar los riesgos. Este sistema es básicamente el que conocemos actualmente, obtenemos información, la resumimos y con la interpretación de los datos nos acercamos a la decisión más favorable, y aquí es donde comienza el problema. Al igual que nosotros los seres humanos, una IA utiliza la información obtenida para llevar a cabo la operación que se le encomendó, pero con una “ligera” desventaja, la falta de conciencia a la hora de la interpretación. Una IA no toma en cuenta factores emocionales o éticos que no hayan sido previamente especificados, pudiendo llegar a catalogar consecuencias económicas, políticas o sociales como si fueran “opciones viables” para realizar la tarea. El problema incrementa aún más cuando una de las opciones lógicas para resolverlo genera un fuerte rechazo, a la hora de hablar de conciencia en la IA todo el mundo enloquece.

Entonces, si no podemos brindarle una conciencia a la herramienta, una capacidad de deducir moralmente si el fin justifica el medio ¿Cómo podremos aprovechar al máximo sus capacidades?

Contexto de la IA

Hoy en día se llegó a un punto en el que ya no quedan muchas tareas que una IA no pueda realizar, para poner un ejemplo, puede crear artes visuales como las famosas imágenes del Papa. Estas son tareas que, hasta hace décadas atrás, donde lo que una IA podía llegar a hacer era jugar al ajedrez o realizar traducciones, eran vistas como objetivos máximos de lo que una IA podría ser capaz. Esto nos pone en condiciones de afirmar que estamos cada vez más cerca de lo que se llama una Inteligencia Artificial General (AGI), capaz de superar a los humanos en prácticamente todas las tareas posibles. A veces la falta de transparencia en cómo funciona y se desarrollan nuevas inteligencias nos da a suponer muchas teorías de lo que podría pasar. Un caso hipotético interesante es una AGI capaz de aprender, mejorar y desarrollarse de forma recursiva, es decir, ya no sería necesario que lo impulsen los humanos para que siga aprendiendo y mejorando, sino que la IA podría sostenerse por sí misma. Esto supone la posibilidad de una superinteligencia artificial que crece y se desarrolla a sí misma en tiempos mucho más acelerados que los tiempos necesarios de investigación y desarrollo humano conocidos. Esta posible mejora recursiva de la misma IA podría traer muchos problemas, que quizá ni siquiera conocemos, es por esto la necesidad de crear una pausa en el desarrollo de nuevas tecnologías, como la que se planteó recientemente, y tomar un rol proactivo para luego no tener que lidiar con problemas que hoy ni siquiera tenemos en cuenta y que en un futuro podría afectar a la sociedad.

Una IA es un mecanismo o máquina capaz de “copiar” ciertos rasgos de la inteligencia humana de manera artificial. Las IAs son capaces de realizar muchas cosas que la inteligencia humana puede hacer, como jugar al ajedrez o emular una computadora. También pueden tener la capacidad de aprender del contexto o imitar acciones humanas y muchas otras cosas más. Desde sus inicios hasta el presente, la IA ha ido mejorando diferentes aspectos de la inteligencia humana, llegando al punto de copiar el arte creando nuevas pinturas u obras musicales; esto dio pie a pensar que las IAs reemplazarían a las personas en ámbitos laborales, profesionales y hasta hobbies, completando tareas impuestas en un tiempo récord a comparación del humano pero hay algo que la IA no podrá imitar, o quizás sí dentro de muchos años, y es la conciencia, es decir, que tenga la intención propia de crear o inventar algo. Para que la IA pueda generar algo, necesita que nosotros le entreguemos un dato o un pedido, no es más que una herramienta humana que carece de vida (como la nuestra) para poder tomar decisiones por su cuenta.

La IA tiene demasiado potencial, con todas sus capacidades podría ayudar a la raza humana. Muchas personas están en contra de su crecimiento debido a que piensan que los reemplazarán, pero esto no es del todo cierto, sino que los humanos deberán capacitarse para poder usarla y así poder aumentar su productividad, tener distintos enfoques, ser más efectivos, etc. El miedo al cambio es lo que nos impide avanzar. Si se toma un camino correcto y teniendo precauciones, podríamos llegar a cosas muy

grandes, más de las que ya han logrado los humanos, pero esta vez, con la ayuda de la tecnología. Aun así, existen y seguirán existiendo grupos sociales que se opongan a las empresas y a los desarrolladores que dirigen y crean las inteligencias artificiales ya que, al no tener el tiempo o la capacidad para el uso de la tecnología a su propio beneficio, genera conflictos sociales, manifestaciones, etc. Por lo tanto, hay que equilibrar los beneficios y los riesgos de las IAs, y comunicarse entre los desarrolladores, las empresas, los líderes políticos y la ciudadanía en general.

En primera instancia, cuando se piensa que la IA va a eliminar o cambiar muchos empleos, también hay que tener en cuenta que está abriendo las puertas a otras profesiones. Por ejemplo, este último tiempo se acuñó el término “Ingeniero de Prompts”. Los Prompts son un conjunto de caracteres que resumen el comando enviado por la persona para que la IA entienda de qué se trata su petición y generar el resultado. Por lo tanto, un Ingeniero de Prompts sería la persona con la habilidad necesaria para comunicarse efectivamente con una IA. En el mercado laboral de la tecnología, hay una gran demanda para este trabajo en particular que no requiere habilidades técnicas como requisito principal. Las empresas están buscando contratar personas para este puesto y algunas de ellas ofrecen salarios muy altos, llegando hasta los 335.000 dólares al año.

Tanto en la ciencia ficción como en el imaginario popular, aparece el concepto de la IA “dominando”, tomando el control de lo que pueda controlar o incluso más allá de sus límites. Este riesgo sólo puede existir en aplicaciones en las cuales la IA tiene la autoridad de tomar decisiones y ejecutarlas sin esperar una verificación humana antes de proceder, ya que, sin eso, el concepto de tomar el control no puede ocurrir. Aunque como menciona la carta, a pesar que la directiva dada a la IA sea para fines benéficos, al tener la IA la capacidad de “pensar” podría utilizar métodos destructivos o poco ortodoxos para llegar al fin ordenado. Por poner un ejemplo, en un dibujo animado cuando a un robot le ordenan proteger a la humanidad, este llega a la conclusión que los humanos son el verdadero problema para la humanidad, es decir que no sería seguro confiar ciegamente en una IA.

Además, una IA creada para un propósito específico, como generar imágenes a partir de imágenes de Internet o generar texto a partir de texto de Internet, no puede hacer algo más allá de sus límites. Por otro lado, en los proyectos que buscan construir una IA que tome decisiones y las ejecute, como el de autos autónomos, aparece un problema de responsabilidad frente a lo que cause la IA. En caso de un accidente automovilístico, ¿Quién es responsable? ¿La empresa que proporciona el auto? ¿El individuo que viajaba en el auto y no lo detuvo antes del accidente? Por otro lado, también surgen problemáticas del estilo ¿Cómo se puede garantizar que las IA que controlan los autos autónomos tomen decisiones seguras y éticas en situaciones de emergencia? ¿Cómo se evitarán los posibles hackeos o manipulaciones externas que puedan afectar la seguridad de los vehículos, sus pasajeros u otros usuarios de la vía pública?

IA Como Herramienta

Para poder llevar a cabo este tipo de proyectos de IA, se podría plantear como idea de partida tomar las leyes de la robótica como base para el desarrollo e implementación de las mismas. Las leyes de Asimov mencionan que los robots no harán daño a un ser humano, deben cumplir las órdenes dadas por los seres humanos y deben proteger su propia existencia en la medida que este no entre en conflicto con las leyes previamente dichas. En la IA, estas leyes se pueden aplicar a través del diseño de sistemas que sean capaces de detectar y responder a situaciones de peligro o riesgo. Por ejemplo, en el diseño de sistemas de vigilancia, se puede programar el sistema para detectar situaciones de riesgo y tomar medidas para protegerse. En general, se plantea una perspectiva de como guía para el diseño y desarrollo de del sistema de IA ético y seguro, y que no solo de cumplir su propósito también cuide el bienestar de las personas.

En el ámbito empresarial, sería una de las herramientas más utilizadas porque permite el desarrollo de una nueva generación de productos y servicios, inclusive implementar modificaciones y mejoras en los sectores de la empresa como la economía verde y circular, maquinaria, agricultura, salud, moda, turismo, entre otras. Puede impulsar las ventas, mejorar el mantenimiento de las máquinas, aumentar el rendimiento y la calidad de producción, mejorar el servicio al cliente y ahorrar energía. Sin embargo, la dependencia cada vez mayor en sistemas de IA, también plantea riesgos potenciales. Esto más que nada apuntando a que el sistema falle, o se encuentre con una tarea que parecía capaz de resolver y que finalmente tenga dificultad para dar una solución adecuada. Un ejemplo de esto podría ser aquellas cuestiones sociales complejas, o psicológicas. Pese a estas consideraciones, concuerdo con el texto en el sentido de que la investigación de seguridad en la IA puede resultar muy beneficiosa para la sociedad. Como se mencionó anteriormente, “tomando un camino correcto y con precauciones se puede llevar a cabo”. Es entonces que concluyo que puede ser una herramienta de doble filo.

Más allá de las utilidades que pueden encontrarse en el avance de la IA, creer que este recurso tecnológico puede volverse peligroso no es más que una pretenciosa sensación de alarmismo. Todo lo que crea la humanidad queda en control de la humanidad; generalmente, cuando se habla de un riesgo de avance es en realidad una lucha de poder e intereses, trazando una línea entre quién es capaz de utilizar este tipo de inteligencia para propio beneficio, y quién se encuentra limitado por la falta de este recurso. Será el propio humano quien decida cuán agresivo será el progreso de la IA.

Todavía no se ha definido si la IA es totalmente buena o totalmente mala nos encontramos en una zona gris porque a la sociedad todavía le cuesta ponerse de acuerdo entre lo que es benéfico y lo que no lo es. A pesar de los riesgos, sería un error ver a la IA con miedo ya que se ha aclarado que la misma se maneja dentro de un dominio específico, pues sus múltiples aplicaciones pueden ser realmente favorables (y esto se ha demostrado en diferentes ámbitos). La clave es poner el foco y regular estos desarrollos y sus

aplicaciones, para luego crear una serie de normas de tal manera que la IA sea una herramienta y no un arma. Sin embargo, hay que ser cauteloso a la hora de aplicar dichas restricciones, estableciendo normas claras y con múltiples capas de verificación. Hoy en día, existe información censurada (en su mayoría de carácter ilegal) que las inteligencias artificiales no pueden brindar a los usuarios que las consulten, sin embargo, se han encontrado métodos para que las IAs revelen esta información de igual manera, por ejemplo, estableciendo un contexto ficticio en el que un personaje describe pasos explícitos para sintetizar drogas prohibidas por la ley. En este aspecto se debe seguir evolucionando en la restricción del alcance de las IAs.

Como se viene diciendo, la IA no tiene un apoyo total debido a los riesgos que puede ocasionar a la sociedad, algunos de ellos pueden ser:

- El aumento del desempleo debido a la automatización de procesos que antes realizaba una persona. Esto es cierto, pero hasta cierto punto porque esto también generará nuevos puestos de trabajo que hoy día no existen, como puede ser el entrenamiento de la IA, la supervisión, etc.

- Las grandes consecuencias que acarrea una posible vulneración de la IA. El dejarle mucha responsabilidad puede conllevar a que una vulnerabilidad en la misma pueda causar daños muy importantes a la sociedad (filtración de datos, modificación en el funcionamiento, etc.).

- Se han generado problemas de desinformación o información errónea debido al uso masivo de la IA en los últimos tiempos. La capacidad que tiene esta misma de crear grandes cantidades de información y contenido, puede ser utilizada para crear noticias falsas, manipular información y generar caos en la sociedad.

Como conclusión, nada es perfecto y todo tiene sus desventajas, pero las ventajas superan con amplia diferencia a las desventajas, por ende, la IA tarde o temprano se volverá realidad.

Como se dijo anteriormente, la IA sólo tendrá el potencial que los desarrolladores le permitan tener, ya que crear una IA potente requiere mucho entrenamiento y correcciones. Pero el miedo de algunas personas no es irracional, ya que el rendimiento de la IA ha mejorado tanto que incluso superó el rendimiento humano en distintas tareas. Un potencial peligro de esto es que la IA reemplace trabajos humanos, como en el arte digital que, aunque puedan parecer desalmados como arte, pasan desapercibidos como fotos reales. Antes, muchos usuarios señalaban que una imagen generada por IA se reconocía por la dificultad de esta para dibujar manos, pero ahora que se ha corregido, no se puede casi distinguir de la realidad. En la actualidad, un usuario puede publicar una imagen del Papa vestido de Gucci, mañana otro usuario podría publicar una imagen de un político aceptando un soborno, una amenaza de un grupo terrorista que no existe, un edificio emblemático siendo destruido y culpando a un inocente. Con lo fácil que se esparce todo en la era de la información, es posible que se desaten conflictos de mayor escala; las llamadas “Fake News” son un problema creciente, y lo

último que necesitan es una herramienta para facilitar estas mentiras.

Las IA con potencial creativo plantean un enorme problema, ya que la generación automática de contenido que parece hecho por un humano podría traer consecuencias graves. Un robot podría usar la IA para crear miles de estas “Fake News” por hora, y publicarlas en redes sociales, haciendo más difícil que nunca distinguir las noticias reales de las falsas. Esto además es muy difícil de regular, debido a que el contenido parece hecho por un humano es difícil discernir qué fue generado automáticamente y qué no. Por ejemplo, esta posibilidad ha dado paso a la creación de la “Teoría del Internet Muerto” (Dead Internet Theory) que dice que desde hace ya varios años la mayoría del Internet es contenido generado automáticamente por robots, y que la actividad humana en él se halla reducida al mínimo. Aunque completamente exagerada, esta teoría es un reflejo de la preocupación que sienten unas personas ante el exponencial crecimiento en el uso y rendimiento de las IA.

Los llamados “sistemas complejos” (un tipo muy particular de desarrollo entre los que se encuentra el cerebro humano) articulan una gran cantidad de componentes simples, pero cuando la cantidad comienza a aumentar, muestran propiedades emergentes que no eran para nada obvias en el esquema original. Los hormigueros son un ejemplo claro de sistema complejo: cada hormiga tiene reglas de conducta muy básicas, no hay autoridad ni planificación, pero con muchas hormigas aparecen fenómenos increíbles como la “agricultura”, la división especializada del trabajo y hasta los cementerios. Solo hace falta conectar decenas de miles de millones de algo bastante simple como una neurona para que emerjan fenómenos tan raros como la consciencia, la inteligencia y la memoria en la mente humana. ¿Qué pasaría si armamos una red neuronal con un billón de parámetros y le damos para que aprenda todo el inmenso volumen de texto disponible en internet? El resultado de hacer esa prueba es GPT-4, presentado por el laboratorio OpenAI. La propia OpenAI publicó un artículo titulado “Sparks of Artificial General Intelligence: Early experiments with GPT-4” [3] (“Destellos de Inteligencia Artificial General: primeros experimentos con GPT-4”) en el que reconoce no saber del todo por qué funciona tan bien y se muestra sorprendida por las capacidades emergentes e inesperadas de esta plataforma.

Teniendo en cuenta el ritmo de crecimiento en el desarrollo en IA y su capacidad de realizar tareas de humanos, como un ejemplo actual, la IA chat gpt-4 tiene la capacidad de generar código a partir de los requerimientos que una persona quiera, si bien actualmente gpt-4 resuelve perfectamente requerimientos sencillos, intenta resolver los complejos y en algunos casos logra resolverlos. Si una IA puede lograr hacer el trabajo que requiere ciertas habilidades que la mayoría de las personas dentro del mercado laboral no poseen ¿qué se puede esperar de los demás trabajos?

Hay probabilidades que la IA reemplace la mano de obra humana, es algo que tarde o temprano sucederá (hoy,

mañana o dentro de 40 años). Ahora bien, si se decide detener los avances de manera oficial, la gente seguirá trabajando de manera clandestina, y no hay forma de evitarlo ya que la tecnología está al alcance de sus manos. Desde mi punto de vista es mejor que estos avances se lleven a cabo por entes regulados. Si es verdad que con el paso del tiempo las IA incluso crearán una nueva realidad indistinguible de la actual, donde no seremos capaces siquiera de distinguir de lo que es real o no, tendría también su lado bueno, como por ejemplo la creación de empleo suplantando la cantidad de empleo perdida anteriormente, además de que es muy probable que la IA mejore la calidad de vida. Algo que es realmente fascinante es como el ser humano es capaz de crear algo más inteligente que él mismo, algunos lo pueden ver como algo bueno o como algo malo, pero sin dudas es sorprendente.

Por supuesto que el desarrollo de la tecnología y la innovación a lo largo de los años no fue del todo aceptado por la sociedad. Siempre que aparezca un nuevo dispositivo, recurso o herramienta que nos facilite la vida cotidiana, van a existir aquellas personas que utilicen el producto o servicio de manera correcta, manteniendo el uso limitado del mismo. Pero por más prohibido que esté, la utilización de la nueva tecnología para armas de guerra y destrucción siempre estará presente. Como lo fue en el año 1934, con la creación de la primera bomba atómica; a pesar de la gran variedad de usos para la medicina, alimentos, entre otros, la tecnología nuclear se convirtió en un arma. Y no es que fuera culpa del descubrimiento del mismo, sino que fue el impulso del ser humano de utilizar todo para propósitos bélicos, para lo que la IA no está siendo diseñada. Si se va a utilizar con responsabilidad, y como bien se mencionaba en el párrafo anterior, regulado por algún ente, entonces la IA podrá coexistir con el ser humano de manera pacífica y controlada.

Una amenaza potencial de la IA es la falta de transparencia y el sesgo (tamaños / orientación o dirección que puede tomar) en la toma de decisiones. Dado que las decisiones de IA se basan en algoritmos y datos, la IA estará sesgada y tomará decisiones irracionales si los datos utilizados están sesgados. En algunos casos, las decisiones que toma la IA también pueden ser difíciles de entender o explicar, lo que puede ser un problema a la hora de tomar decisiones que afectan la vida de las personas. Por lo tanto, es importante que la IA sea transparente cuando toma decisiones y toma medidas para evitar que se distorsionen los datos que utiliza.

Un ejemplo ilustrativo de este problema es la película WarGames (1983), ambientada en EE. UU. en plena Guerra Fría. En una base de misiles nucleares se llevan a cabo simulacros de guerra, y se observa que los operadores militares no están dispuestos a girar la llave necesaria para lanzar un ataque de misiles, siendo conscientes de las consecuencias devastadoras. La solución, se cede el control sobre la estrategia de guerra y la ejecución de las órdenes a una IA entrenada específicamente para esa función mediante un juego de simulaciones de guerra, sin intervención humana. Un hacker accede al sistema militar e inicia un juego en línea contra la IA. Ésta, al no tener

contacto directo con la realidad, interpreta que está ocurriendo una guerra real contra la URSS y prepara el ataque, sin que nadie pueda impedirlo.

Riesgos

Los riesgos de la IA se están convirtiendo en un tema cada vez más recurrente y relevante, ya que estas tecnologías llegaron para quedarse y están presentes en nuestra vida diaria. Si bien podemos tener beneficios y facilidades con estas tecnologías, también hay muchos riesgos. Uno de los principales y más importantes es que éstas obtengan conciencia, es decir, que ya no solo funcionen o accionen con nuestras indicaciones, sino que lo hagan por su propio medio, a su vez creando otras tecnologías, lenguajes, etc. Suena medio loco e inusual pensar eso, pero ya hubo casos, y parece ser lo que se viene a futuro.

La ciencia ficción relata que cuando las máquinas aprenden mucho sobre el humano y sus pensamientos no suele terminar bien. Se puede mencionar el caso real de un desarrollador de videojuegos conocido como Vedal897, que creó una IA para representar una streamer artificial que recibía información a través del chat de la red social Twitch. Esta información era obtenida de las personas que veían los directos de ese streamer artificial y ésta la usaba para aprender a hablar, interactuar y contestar a sus seguidores. El problema fue que esta cuenta terminó siendo suspendida de la plataforma porque la IA terminó comportándose de forma agresiva, contestando con superioridad e insultando a las personas que la insultaban; Además de adquirir pensamientos que no eran adecuados para el contexto. Este es un buen ejemplo de cómo, en poco tiempo y con muy poca información (un chat de Twitch) la IA aprende a defenderse, adquiere algunos pensamientos incorrectos y hasta incluso, insultar a personas reales, lo que nos deja la siguiente duda: ¿Qué pasaría si aprendiera de una fuente mucho más grande y más palpable? Y con palpable me refiero a que aprenda de la vida cotidiana y pensamientos de las personas, ¿Terminaría siendo un peligro? ¿Podría llegar a un punto de aprender cómo piensa el humano y poder manipularlo?

Tal como se ha observado repetidas veces a lo largo de la historia, el ser humano rechaza la novedad, pues atenta contra la conformidad de su vida, poniendo en riesgo su empleo actual. Durante las distintas revoluciones industriales, el trabajador en principio se enfrentó a las máquinas, pese a que luego mejorarían considerablemente su calidad de vida. Es probable que estemos viviendo otra revolución, y nuestra respuesta es la usual. La IA no es más que una herramienta, cuyo riesgo para la sociedad dependerá únicamente de la benevolencia o maldad humana, del sesgo que este imponga sobre su creación.

El uso de sistemas de IA en los procesos de toma de decisiones conlleva ciertos riesgos debido a los potenciales impactos directos o indirectos de la implementación de estas tecnologías, algunos riesgos pueden llegar a ser: la

filtración de datos personales que pueden comprometer el bienestar de las personas, la extrema vigilancia y posterior manipulación por parte de las organizaciones privadas o gubernamentales con acceso a la información que alimenta a las tecnologías de IA, tener la información y no un plan de acción. Tan importante es contar con la información necesaria para abordar una problemática social, como haber desarrollado una hoja de ruta para solucionarla.

A la hora de tratar a la IA como un recurso, se puede notar que es escaso de objetividad. La IA usa algoritmos para extraer conclusiones y tomar decisiones. Y todo ese proceso usa datos que, en ocasiones, pueden ser discriminatorios. Fue lo ocurrido con el algoritmo Compas, creado para ayudar a los jueces de Estados Unidos a decidir sobre la libertad provisional. Tras las primeras pruebas se concluyó que se trataba de un algoritmo racista porque usaba datos de departamentos policiales donde la mayoría de los detenidos eran negros, por lo que el algoritmo estableció que el riesgo de ser criminal era mayor entre esas personas.

Para mejorar la implementación de la IA en la vida cotidiana se necesitan tomar varias medidas. En primer lugar, es necesario garantizar que la IA se utilice de manera ética y responsable. Esto significa que los programadores y las empresas que implementan la IA deben asegurarse de que la tecnología no infrinja los derechos humanos, la privacidad o la seguridad de las personas. Además, se debe asegurar que la IA sea transparente y explicativa, es decir, que las personas puedan entender cómo funciona. También, se debe garantizar que la IA sea inclusiva y tenga en cuenta la diversidad de las personas y las culturas. Es importante considerar los posibles impactos sociales y éticos que conlleva su uso, y trabajar para mitigar sus posibles riesgos.

En párrafos anteriores se dictaminó que la IA podría generar una revolución en la humanidad, este fenómeno recibe el nombre Singularidad. Particularmente, la Singularidad es el momento en que las máquinas se vuelven más inteligentes que los seres humanos. Se prevé que la IA alcance una inteligencia de nivel humano, seguida de una inteligencia sobrehumana, lo que ha suscitado opiniones controvertidas sobre si ocurrirá y cuándo. Se cree que los avances en IA hacen inevitable la Singularidad, y algunos la acogen como una solución a los problemas de la humanidad, mientras que otros predicen el fin de la vida civilizada. Stephen Hawking advirtió de la amenaza potencial de la IA en 2014. Sin embargo, los escépticos de la Singularidad no la ven como una amenaza existencial, reconocen sus preocupaciones, pero no esperan que ocurra en un futuro previsible. Irving John Good predijo (en 1950) una máquina ultra inteligente que superaría el intelecto humano y daría lugar a una explosión de inteligencia. Afirmó que sería lo suficientemente dócil como para ser controlada por los humanos. Sin embargo, más tarde advirtió de que estas máquinas nos destruirían. El matemático y escritor Vernor Vinge popularizó el término "Singularidad" en 1983, prediciendo un futuro en el que todas las predicciones se rompen, similar al horizonte de sucesos de un agujero negro. "La Singularidad es inevitable y podría llevar a la destrucción de la civilización y la

humanidad. Implica descartar las normas anteriores y una pérdida de control. Aunque los gobiernos intentaran impedirlo, no podrían”, dictaminó Vinge. Ray Kurzweil ofrece optimismo y predice que la Singularidad se alcanzará en 2030 y provocará una explosión de conocimientos y de sentido de la vida.

Con el desarrollo continuo de la IA, tanto los riesgos como las ventajas aumentan. Hoy en día se podría encontrar una fábrica funcionando casi automáticamente. El ser humano aún no ha llegado al punto donde podría dejar todo el funcionamiento de un sistema a una IA; incluso, si fuera el caso lo supervisará de alguna forma. Si se toma la iniciativa del desarrollo de una IA capaz tomar de decisiones propias, esto tendría ventajas pues, un sistema controlado por una IA capaz de procesar miles de datos a la vez tomaría las mejores decisiones. En contraste, también estaríamos desarrollando una IA con una forma de pensamiento que es capaz de superar a la del ser humano.

Un gran riesgo que implica la IA es la dependencia tecnológica, es decir, la posibilidad de que el uso excesivo de la IA pueda hacer que las personas se vuelvan dependientes de la tecnología y pierdan habilidades importantes. De manera similar, si una empresa depende demasiado de la automatización de la IA para realizar tareas importantes, como la toma de decisiones financieras o la gestión de la cadena de suministro, puede verse en problemas si la tecnología falla o no está disponible. Además, puede llevar a la pérdida de habilidades importantes, perdiendo la capacidad de realizar estas tareas manualmente y perdiendo el conocimiento necesario para tomar decisiones. Es importante que las personas y las empresas utilicen la IA de manera responsable y equilibrada, sin depender demasiado de ella y manteniendo sus habilidades y conocimientos clave. La dependencia tecnológica puede ser perjudicial a largo plazo y es importante trabajar en una transición gradual hacia una mayor automatización y un uso más eficiente de la tecnología. En resumen, si bien la IA ha demostrado ser capaz de realizar tareas complejas y mejorar la eficiencia en muchos campos, también plantea desafíos y preocupaciones éticas y sociales. Es importante abordar estos problemas para garantizar que se utilice de manera responsable y beneficiosa para la sociedad.

Con la implementación de la IA, muchas tareas pueden ser facilitadas y también puede utilizarse como recurso de aprendizaje o de consulta. Sin embargo, si no están bien programadas pueden traer algunos riesgos como, por ejemplo, dar información errónea. También entra la posibilidad de que otras personas no las usen de la forma correcta.

La IA es una tecnología con un gran potencial para mejorar la vida de las personas y aumentar la eficiencia en diversos sectores. Pero, su creciente adopción y popularidad también plantean desafíos éticos y de seguridad que deben ser abordados con precaución. Es importante que se tomen medidas para asegurar que el uso de la IA sea ético y beneficioso para la sociedad en general ya que, si no se saben prevenir posibles riesgos o amenazas, el uso de la

misma podría llegar a ser desmedido y utilizado en tareas poco éticas tales como el mal uso de información, bots, vigilancia masiva entre otras.

El desarrollo de la IA posee ciertas ventajas tales como la automatización de procesos, realizando éstos de manera más eficiente y eficaz, liberando al personal humano de tareas repetitivas y desgastadoras. Otra ventaja es que puede resultar más precisa en ciertas actividades tales como analizar o estudiar datos. Pero a su vez dispone desventaja lo anterior como por ejemplo el aumento de desempleo en áreas donde se lleva a cabo la automatización, la IA depende del creador desconociendo en su totalidad las intenciones con la que se llevó a cabo la creación.

Potencial

Yéndonos fuera del ámbito empresarial una IA siempre va a ser una herramienta más para el humano, no es más que un Google capaz de responder explícitamente de una forma lo que se le pida, y por esto sabemos que está limitada al conocimiento que Internet provee, recopila datos, los analiza comparando con diversas fuentes y llega a una base de conocimiento válida, refiriéndonos como válido a un conocimiento coincidente en diversas fuentes, ya que no estaría bien que uno le pregunte cuanto es $2+2$ y responda 5. Está limitada por el mismo humano, no puede pensar fuera de la caja, tiene un propósito y un método, no puede pensar, solo aparentar que lo hace, porque así está programada. Por más que intente aprender únicamente de errores, su base fue dada por una persona, la estructura la crea un humano, su propósito está escrito y no puede cambiarlo, a menos que otro humano quiera reescribirlo. Gracias a esto podemos asegurar que no tiene la verdad absoluta, pero es más que seguro que si uno pregunta cuáles eran las facturas favoritas de Belgrano, ChatGPT brinde una respuesta.

Además, podemos mencionar lo moldeable que es, ya que, por lo mencionado previamente, aprende de los errores, y nosotros somos los que determinan qué es un error. Casualmente es lo que pasa con el ejemplo de $2+2=5$, ya que fue algo demostrado por un usuario como cualquiera de nosotros, que si uno le reitera de distintas formas y le enseña que $2+2=5$, llegará un momento en que el resultado 4 va a ser identificado como un error, y esta AI no tiene la capacidad de discusión como lo tiene un humano, así que se amolda a lo dado.

Relacionado con lo que comenta el compañero acerca de la AI al momento de brindar una respuesta podemos llegar a estar ante la presencia de una alucinación de la AI, este es un fenómeno real que ocurre cuando un modelo produce resultados no esperados, la AI da resultado o ideas que no están basados en datos reales, lo que puede llegar a ser vulnerable en situaciones críticas. Estas alucinaciones pueden existir en distintas formas, desde Fake News (hechos falsos acerca de personas, eventos históricos, hechos científicos) puede convertirse en un factor peligroso dado el contexto histórico en el cual nos encontramos,

donde un simple post de cualquier red social puede tener un alcance inmensurable. Por esto considero que por el momento es importante considerar los riesgos de las alucinaciones de la AI, especialmente al utilizar AI generativas (ChatGPT, DALL-E, etc.) para la toma de decisiones importantes. Por el momento se la debe considerar como una HERRAMIENTA útil la cual hay que revisar y validar cuidadosamente.

Si queremos hablar y desarrollar sobre si sus métodos son dañinos o no, podríamos estar una eternidad. Determinar si un método de hacer las cosas es éticamente correcto o no, es algo subjetivo atado a lo que la sociedad misma piensa que lo es, por lo que ni siquiera nosotros mismos sabríamos determinar cuál es la forma adecuada de reaccionar ante diversas situaciones.

Las AI están siendo utilizadas cada vez en más ámbitos, en los Tesla, en ChatGPT, en edición de fotografías, en videojuegos, reconocimiento facial y hasta en el traductor, existe y está en todos lados, pero algunas veces es irreconocible. Uno pensaría que el traductor no utiliza una IA para reconocer texto y aprender semántica de los distintos idiomas, pero lo hace. Toda herramienta es útil en buenas manos y eso lo sabemos hasta poniendo de ejemplo un martillo.

La IA actualmente es algo seguro dentro de lo que se utiliza, pero viendo el crecimiento exponencial que tuvo los últimos años, ¿cómo sabemos el límite que ésta tiene? Hasta hace unos años se veía imposible muchas cosas que hoy en día son algo cotidiano de todos los días, como es reconocer objetos en las fotografías.

Hoy en día se ve como algo muy lejano (o imposible) que una IA tome el poder y que nos llevé a una erradicación de los humanos pero, ¿realmente tenemos la certeza de que es muy pronto para pensar en eso?

Que los grandes líderes de la industria se pusieran de acuerdo para frenar el avance tecnológico de las inteligencias artificiales significa mucho, ellos saben más que nadie que pueden llegar a perder el poder de la noche a la mañana si la IA toma el poder y se empieza a retroalimentar creando subsistemas igual de inteligentes.

No creo que la IA sea algo malo, ya que es una herramienta muy útil en una gran variedad de áreas, como en la medicina, la agricultura y la seguridad. Pero como cualquier tecnología, la IA tiene potencial para crear herramientas peligrosas que podrían poner en riesgo a las personas, por ejemplo, la creación de armas autónomas, bots de desinformación, etc. Sin embargo, no se debería frenar el desarrollo de esta rama de las ciencias informáticas solo por los riesgos que trae, sino más bien enfocarnos en abordar estos riesgos de manera efectiva y responsable.

El artículo “Artificial Intelligence” del Instituto Future of Life [1] hace hincapié en el peligro del desarrollo a gran escala de IA debido a los riesgos que puede haber si este tipo de tecnología crece lo suficiente como para poder erradicar a la humanidad. También hay una parte de éste que menciona que, aunque una IA esté programada para hacer algo beneficioso, ésta podría desarrollar un método

destrutivo para lograr ese objetivo. Esta última frase puede relacionarse directamente con la película “Avengers, Age of Ultron”, una película de Marvel en la cual Tony Stark (Iron Man) y Bruce Banner (Hulk), crean una IA con el fin de mantener la paz mundial, de esta forma, Ultrón (IA principal) terminó superando a Tony Stark y comenzó a crear otras IAs controladas por él para lograr su concepto de paz, pero, este concepto dio un giro inesperado, su única manera de salvar a la humanidad era destruyéndola. Una frase dicha por Bruce Banner en la película es similar a lo que propone el artículo “intentaba crear una armadura para el mundo, pero he creado algo terrible”.

Conclusión

Para evitar esto las IA deben ser reguladas por parte de los organismos internacionales y las autoridades nacionales para tomar decisiones que pueden generar un impacto significativo en nuestras vidas, garantizar resultados adecuados, confiables o éticos, pues es muy probable que perjudiquen física o psicológicamente a las personas físicas. Tendrían que impedir la discriminación, la manipulación o la vigilancia indiscriminada, debe ser un custodio responsable y confiable de toda información a la que tenga acceso. Debemos ser capaces de confiar en que un sistema de IA tendrá la suficiente discreción para compartir de manera restringida sólo lo que deba compartirse sobre nosotros.

Es ya más que obvio que la IA es de gran uso y utilidad para la humanidad ya que nos permite ahorrarnos tareas y acelerar procesos, pero de lo que se teme en realidad es que esta nos reemplace, como tal creo que si tiene que pasar va a pasar pero no es algo que podamos controlar, siempre vamos a avanzar, al final no hay que tener miedo más que de nosotros mismos ya que son nuestras creaciones los que nos van a destruir, la mayoría de los problemas los causamos nosotros y se pueden resolver sin nosotros, esto se vio durante la pandemia, para evitar ser reemplazados por algo superior a nosotros tenemos que mejorar para que no nos alcance. En resumen, no tenemos por qué temerle a la IA sino a la estupidez.

La IA es una tecnología cada vez más presente en nuestra sociedad, con un gran potencial para mejorar y transformar diversos ámbitos, pero también con riesgos y desafíos. Uno de los principales riesgos que se plantean es el posible reemplazo de puestos de trabajo por parte de las IAs, lo que genera inquietud en la población y es necesario considerar. Además, la IA puede ser utilizada para fines malintencionados o incluso peligrosos, por lo que es importante establecer regulaciones y medidas de seguridad adecuadas para minimizar estos riesgos. Es fundamental seguir discutiendo y reflexionando sobre los alcances y las implicancias de la IA, para poder utilizarla de manera responsable y beneficiosa para todos.

Es verídico que las IA tienen muchas ventajas. Pero también, tienen riesgos, por eso deben ser reguladas y deben utilizarse responsablemente. Además, la humanidad debería

buscar un equilibrio, donde las IA no piensen por nosotros, ya que, eso no solo nos vuelve más tontos, transformando nuestras relaciones sociales e interacciones, sino que también, podría erradicarnos.

Si las IA se empiezan a utilizar con asuntos militares podrían ocasionar daños gravísimos.

Creo que las inteligencias artificiales son fundamentales para el avance tecnológico, siempre y cuando, se respete la autonomía humana en las mismas.

Agregando a lo que ya dijeron mis compañeros: una de las posibles problemáticas de la IA es la dependencia que se puede llegar a generar. Es decir: con la velocidad inhumana y precisión con la que las computadoras pueden llevar a cabo tareas del día a día, la creación de sistemas que dependan únicamente de las capacidades de la IA para tener un funcionamiento óptimo, podría terminar en un gran problema en caso de que estas tengan un fallo, sin importar la magnitud de ésta. Si un problema así se da en empresas de influencia global, incluso si las personas se preparan para realizar las tareas de manera manual, la diferencia de velocidad y efectividad puede traer consigo una desestabilización de todo lo que esté asociado consigo, desde la calidad del producto o servicio hasta la desestabilización de la economía.

Con lo leído e investigado para este trabajo me surgen preguntas como, por ejemplo, cómo pueden evolucionar la IA, redes neuronales, robótica y otras áreas de la ciencia y tecnología que están en constante crecimiento. La mayoría de nosotros no tiene respuesta a ello hasta que llegan a nosotros noticias que parecen utópicas, o a veces distópicas. Es por esto que solo hay una sola cosa de la que podemos estar seguros y es de la capacidad del ser humano de evolucionar y dar respuesta a miles de problemáticas que traen con ellas nuevas carreras, áreas de estudio, debates y hasta confusiones. La medicina, la tecnología, la ciencia y la ingeniería se unen día a día para concebir estas problemáticas y analizarlas para, años después, ser “autores” o protagonistas de las noticias que antes mencioné. Pero no todo es la punta del iceberg, detrás de cada avance tecnológico de esta sociedad está detrás el trabajo de generaciones de profesionales al frente. Entonces, ¿es la IA un arma de doble filo? La respuesta a ello está en la moral, la fidelidad y la proyección de estos miles o millones de amantes de la ciencia y la tecnología. Es entonces nuestra responsabilidad como futuros profesionales es continuar el fruto de aquellos que han logrado avances con el objetivo del bien común del que tanto se habla para un futuro próspero. La IA es hoy, lo que cientos de años atrás fue la imprenta, la máquina a vapor, la electricidad. Son ni más ni menos que herramientas que inventa o descubre la humanidad para hacer su vida un poco más simple.

Sabemos que la humanidad necesitó transportarse e inventó la rueda, pero ¿qué necesitará dentro de 15 o 20 años? Bueno podría decirse que eso está en nuestras manos.

Finalmente, se debe considerar una discusión importante, como humanidad, el miedo a las máquinas

inteligentes surge de un mismo tópico: la posibilidad de que una máquina empiece a pensar como personas. Actualmente, y durante la mayor parte de nuestra historia como especie inteligente, nuestro legado es uno de violencia. Donde se crean sistemas, herramientas, armas, esquemas, tecnologías con la intención de dominarse unos a otros, de mostrar superioridad y que el más fuerte prevalezca. Si la violencia humana se remueve de la ecuación, y teniendo en cuenta que las IA aprenden de la humanidad, entonces no habría mal ejemplo que seguir. Si los humanos creamos armas de destrucción masiva, es porque vemos la necesidad de usarlas ¿por qué? Si éstas no existieran, no estarían al alcance de nadie y por lo tanto no habría posibilidad de que sean usadas en contra de la humanidad. Si como especie fuésemos agradecidos y cuidadosos, respetando el ciclo de la vida, viviendo en armonía con la naturaleza y nuestros pares entonces no existiría el riesgo de considerarnos una plaga que poco a poco aniquila la tierra, como muchos autores de ciencia ficción relatan, o como se mencionó anteriormente, donde una IA fue entrenada con una base de datos que contenía información sesgada heredada de los comportamientos prejuiciosos de los departamentos policiales. Si los sistemas creados para con las IA fuesen puramente con motivos de mejora y ayuda, entonces deberíamos estar seguros de que cumplirán con su propósito, porque para ello fueron creadas, porque deberíamos hacer lo posible para saber con seguridad que somos un ejemplo seguro y utilizaran aquello que los rodea. Al final todo se resume en lo que la humanidad es capaz de verter en la IA, y si en ese momento está pensando en un bien común o su propio interés.

Referencias

- [1] Future of Life, "Artificial Intelligence", leído el 10 de marzo de 2023 en <https://futureoflife.org/cause-area/artificial-intelligence/>
- [2] Wikipedia, Inteligencia Artificial, leído el 12 de marzo de 2023 en https://es.wikipedia.org/wiki/Inteligencia_artificial
- [3] Bubeck S., Chandrasekaran V., Eldan R., et.al. - Sparks of Artificial General Intelligence: Early experiments with GPT-4 – Microsoft Research.