TRANSCRIPCIÓN AUDIO

F: Hola Jorge, ¿qué tal? muchas gracias por esta charla.

J: Por favor, al contrario. El gusto es mío Fabio.

F: Jorge, para empezar a hablar tal vez sería interesante que empecemos con el dato y pensar que el dato que está como en la raíz del big data, de los algoritmos, si el dato es algo objetivo, el dato es algo que podemos dar por sentado o el dato tiene sesgos.

J: El sesgo que se tiene es en la recolección del conjunto de datos, que hay un sesgo humano, y ese sesgo se traslada después al modelo, uno habla de un modelo como algo neutro sin sesgo, y esa idea es prácticamente imposible desde el momento que los datos y cómo uno obtiene esos datos, los selecciona, los prepara, ya tiene también un sesgo.

F: Se hablaba del problema del dato, que parece siempre que el dato es objetivo pero el dato trae en sí mismo un conflicto. Digamos, la altura del río tal vez no, hay datos menos cuestionables pero, en educación sobre todo, cuando uno dice ¿qué significa sacarse un 10?, o sea ¿cómo medir el éxito en educación? Existe la calificación, pero podés discutir si la calificación es un dato válido para decir que alguien tiene éxito

J: Claro, exacto, además esos datos es una acumulación de datos generados con el sesgo de cada humano. La escala de calificación puede generar un sesgo distinto de cada humano y hay muchas cosas, por ejemplo sistema de predicción de delitos, se puede dar en un lugar que tenga una fuerte correlación con una raza, con un grupo étnico, por ejemplo en EEUU porque uno ve que un alto porcentaje de la población carcelaria pertenece a un grupo racial determinado y entonces uno puede inferir erróneamente que la persona por su color de piel tiene más tendencia a cometer un delito cuando hay otros factores que hicieron que haya una mayor cantidad de personas de determinada característica étnica en la cárcel, condiciones sociales, prejuicios, prejuicios de los jurados, de los jueces y todo ese prejuicio, todo ese sesgo está en los datos, entonces se traslada al modelo. Por eso cuando decimos, las máquinas están libres de sesgo, no, al contrario, tienen toda la herencia. Ese es un punto muy interesante que tratamos que es uno de los grandes problemas de la IA sobre todo cuando las vamos a aplicar en lo social, porque ¿qué terminamos haciendo?, trasladamos a los datos nuestros prejuicios, entonces ¿qué vamos a terminar diciendo? así como yo cruzo de vereda cuando veo al *morocho*, vamos a hacerlo así cuando uno ve una persona así porque la televisión te bombardea, y cada vez que hay un caso de inseguridad, terminan metiendo a dos *pibes morochos* de gorrita en un patrullero, entonces te termina quedando eso en la cabeza, si

un modelo es alimentado con datos similares, va a terminar teniendo ese prejuicio o ese sesgo.

F: La profecía autocumplida

J: Claro, es lo que pasó también con unos "bots", un experimento que se hizo hace un par de años, lo agarró un par de chicos que *bulleaban* a otros y el robot terminó diciendo frases racistas que eran los sesgos del ambiente, entonces si vamos a empezar a confiar como si fuera una verdad superior la IA, hay que tener cuidado con esa alimentación con sesgo. Ese es un tema.

F: Cuando vos hablás del análisis de datos, ¿hasta dónde vos empleás fuerza de la máquina pensando los criterios que vos le diste, humano, vos armaste el algoritmo y en dónde empieza la máquina a conformar el propio algoritmo? osea, ¿en dónde empieza la IA y en dónde las definiciones humanas?

J: Bien, la IA es un término muy difuso, en realidad la parte artificial está clara, no generada por humanos. La parte de inteligencia es muy difusa, en general es más fácil, no vas a encontrar definición única de IA, más allá de lo genérico que consiste en percibir el entorno y generar decisiones propias, en general hay ramas dentro de la IA que sí tienen definiciones más concretas, la más utilizada y en este caso que se utiliza mucho en la ciencia de datos es lo que se llama aprendizaje automático, aprendizaje automático es se dice que un algoritmo tiene aprendizaje automático cuando mejora su eficiencia al realizar una tarea a medida que aumenta su entrenamiento, el entrenamiento es conocer más datos, pero el algoritmo no se toca, a diferencia del algoritmo tradicional, un algoritmo tradicional, la programación tradicional un algoritmo va a funcionar igual desde el día 1 hasta el día un millón.

F: Lo cargó un humano el algoritmo

J: Tiene reglas fijas.

F: Tiene reglas fijas y la programación típica que uno le dice a la máquina lo que hace y la máquina responde.

J: Eso es la programación tradicional, no tiene en cuenta el historial y la experiencia. El algoritmo de aprendizaje automático analiza el historial y va corrigiendo, va creando sus propias reglas, no viene un humano y le dice este producto lo vas a enviar a la gente que tiene menos de 21, a los de 45, se va dando cuenta en base al historial y se va corrigiendo. Si cambia la tendencia, va a ajustar la regla, si este producto lo deja de comprar la gente de 21 y lo empieza a comprar la gente más grande, va a empezar a darse cuenta y va a reconocerlo, va a mejorar su eficiencia y tiene que mejorar su performance de productos vendidos, no sobre cantidad de recomendación, uno también tiene que poder medir eficiencia. Eso es el

aprendizaje automático. La IA en general la usamos para problemas donde no hay certidumbre, es decir uno utiliza programación clásica para caminos directos “si es A, si es B, si es A va por un lado, si es B va por el otro”. La IA la usamos para programas donde no tenemos certeza ¿quién va a comprar un producto? no es un problema de certeza, una apuesta deportiva: dame la probabilidad ¿quién va a ganar el domingo tal partido? ¿cuál tiene más probabilidad? la probabilidad del embarazo que hablábamos, son problemas que no es una clasificación automática, no es un robot, bueno, vas a pintar tal modelo de auto, tiene la puerta tal dimensión, ya queda programado eso y fijo. Si viene otro auto tenés que reprogramar con otra dimensión y tenés que reprogramar al robot con las dimensiones de ese auto. Esto no, la IA es como si yo te dijera bueno programo un brazo robótico para que pinte cualquier puerta de auto que le traiga, aunque sea una puerta nueva que nunca haya pintado, que pueda reconocer que es una puerta, dónde están los límites y pintarla sin dejar ningún lugar sin pintar, eso sería la diferencia entre un programa tradicional y un programa con IA, percibir el entorno y aplicar nuevas reglas, que una puerta nueva con dimensiones que nunca pinté, la pinto toda. Eso más o menos sería.

Y con respecto a aprendizaje automático siempre hago una analogía, el entrenamiento de una máquina es la cantidad de datos que recibe, datos históricos y lo que es la experiencia del ser humano, es muy manifiesta ahí la similitud. Es decir, así como las personas mejoramos o adquirimos mayor expertiz o sabiduría con la experiencia, las máquinas análogamente lo hacen con el entrenamiento y con los datos. Cuando una máquina tiene poco entrenamiento, pasa lo mismo que cuando un humano tiene poca experiencia, tiene mucho sesgo y tiene prejuicio, es como cuando uno va por la calle y cruza de vereda porque vio alguien que no le gustó, ahí actuó el prejuicio, vos actuaste a costa del prejuicio porque por ahí te robaron una vez en la vida y era una persona con determinadas características físicas entonces cada vez que ves una persona similar, cruzas de vereda. Eso es un prejuicio, digamos, uno tiene un sesgo alto, tiene poca varianza, porque son pocos casos, entonces uno no es un buen identificador, un policía por ejemplo, seguramente es mucho mejor que uno en detectar personas sospechosas o no, te puede decir, no, mirá. Porque tienen mucha más experiencia, además uno está muy afectado por su experiencia personal, si a uno lo asaltaron dos veces, te queda eso, es muy fuerte, y lo mismo pasa con la máquina, si entrenaste con tres ventas y tenés un historial de tres ventas de café, y te compraron dos personas de 30 años y una de 15, le vas a ofrecer a una persona de 30, tal vez alguien de 15 y si un tipo tiene 50 no le vas a ofrecer nunca un café.

F: Cuando vos hablás de aprendizaje automático hablás de machine learning.

J: Exactamente, es eso, aprendizaje automático.

F: En todos los esquemas aparece siempre el mismo dibujito. Es el de las tres hileras, los puntos que representan neuronas entrelazadas.

J: En realidad eso que están viendo son redes neuronales, que es un algoritmo, es una parte de aprendizaje automático. Aprendizaje automático es una amplia gama de algoritmo, incluso de tipos de aprendizajes automáticos, está el supervisado, el no supervisado, el semi supervisado o refuerzo.

F: Igual, a lo que voy, no es nuevo...

J: No, no, por supuesto que no es nuevo

F: Y yo recuerdo la detección de cara de Picasa del software de Google, que debe ser del 2009, cuando empezó a reconocer caras, van a ser casi 10 años. Y vos lo entrenabas, porque vos le ibas diciendo el nombre de las personas y después el software te proponía esta persona es ésta y vos le decías sí, no, sí, no y después se va entrenando y reconoce cada vez mejor.

J: Sí, ahí habría que ver si es un tema de IA o de almacenamiento de parámetros porque a lo mejor por ahí ya maneja un nivel de certidumbre al nivel de una distancia entre ojos y es como una búsqueda en una base de datos más que un algoritmo, pero tendría que ver más ese algoritmo. Mucha gente por ejemplo habla del ajedrez como un caso de IA, y no, en general los programas de ajedrez son programas de búsquedas con heurísticas y poder de cálculo, por ejemplo, el programa que derrotó Kasparov hace 20 años, deepBlue, era un programa con fuerza de cálculo y heurísticas, no era un programa que después de jugar un millón de partidas juega mejor que la partida uno, no, juega exactamente igual, no es un programa que mejora con entrenamiento.

F: porque el ajedrez lo usan y Kasparov tiene su programa a partir de la IA y su programa... ¿y lo del GO también es lo mismo?.

J: No, lo del GO no, lo del GO es un ejemplo de aprendizaje por refuerzo, ese sí es un programa clásico de aprendizaje automático y después sí salió un programa de ajedrez competente, realmente muy bueno, que se llama Alpha Zero, hecho a partir de Alpha dos, es un programa de ajedrez sí con IA y aprendizaje automático. Pero los programas clásicos de ajedrez o los que uno tiene en la computadora no es un ejemplo de IA, o por ejemplo de aprendizaje automático, es un tema de heurística y búsquedas, es más difuso. Se lo trata más como fuerza bruta.

F: Subyace esta idea de que la computadora puede calcular, buscar rápido, por eso decimos fuerza bruta, pero su forma de trabajar es siempre la misma, no cambia, no evoluciona con el tiempo, y en el otro caso, con el aprendizaje va cambiando en la medida que va evolucionando.

J: Exacto, un programa que empieza a entrenarse, apenas sabe las reglas y juega como un principiante y a medida que va jugando partidas va mejorando su fuerza, se va premiando y castigando en función del resultado y se va ajustando, entonces va diciendo "no, por este camino me fue mal" tiene memoria, "cuando fui por acá, me fue mal, perdí", prueba una nueva jugada y empieza a inferir reglas, más importante, no es que memoriza el camino sino que empieza a inferir reglas. Porque ése es el punto de la IA, que es a partir de la experiencia, no

es que tengas memorizado el camino, porque si no no te serviría para un caso nuevo, sino que puedas inferir reglas para enfrentarte casos nuevos, y para eso es que se usa la IA, hacer programas que puedan interpretar el entorno y aplicar y deducir reglas para trabajar o para tomar acciones ante casos nuevos. Como cuando uno estudia, el profesor te da ejercicios, vos los hacés, pero en la prueba te da un ejercicio nuevo, supuestamente te da un ejercicio que nunca viste, pero tenés que estar capacitado para resolverlo, es así cómo te evalúa, eso se hace de la misma forma con la IA. A un programa tradicional uno podría enseñarle ejercicios pero le das un ejercicio nuevo y no sabría resolverlo.

F: Y ahora, en el aprendizaje automático de las máquinas, ¿las máquinas aprenden una porción chiquita específica?

J: Son muy específicas, estamos muy lejos de algo poco humeante, que se llama la IA general, sería como algo más Terminator, o más un Hall 9 mil por ejemplo, que puede interpretar una conversación, y que pueda aplicar IA a múltiples propósitos. Por ahora, no, vos tenés un sistema que juega al GO y solo sirve para el GO, vos lo querés hacer para el ajedrez, tenés que reprogramarlo, etcétera. Vos tenés un sistema de recomendaciones, bueno, sirve para recomendar, después bueno, va a empezar a entrecruzar sistemas inteligentes. Pero la IA general todavía es algo muy lejano. Además, dentro de IA así como está aprendizaje automático, están otros temas como son los los agentes inteligentes autónomos que también tienen aprendizaje automático pero tienen también el tema de la percepción del entorno.

F: El diseño por ejemplo de un coche.

J: Exacto, lo que se llama por ejemplo computer vision, todo lo que es internet of things (Internet de las cosas) con detectores, con sensores y después tenés algo todavía más complejo que es el procesamiento del lenguaje natural, que es de lo más complejo en la IA, que es interpretar la conversación y dar una respuesta semánticamente válida y también simbólicamente válida, con contenido válido, eso también es algo muy árduo.

F: Todavía falta.

J: Sí, sí, sí, eso está experimental, en inglés están más avanzados, tienen los dispositivos del estilo de Alexa.

F: Me acuerdo del caso este que fue el año pasado de las computadoras que empezaron a hablar, y saltaron del inglés y generaron su propio lenguaje.

J: Sí, empezaron a formar el lenguaje, a repetir tres veces los verbos, sí, ahí estaban entrenando, es entrenamiento justamente, empieza a escuchar al otro, y en lugar de escuchar al humano, empieza a escuchar a otra máquina.

F: De hecho armaron su propio código.

J: Claro

F: Lo venden como que fue un fracaso.

J: Yo no diría que fracaso, fue colateral, a mi me pareció muy interesante.

F: Porque ahí evolucionó el lenguaje, el lenguaje evoluciona no? Lo que pasa es que a nivel social el ritmo de evolución es más lento Cuando eran tres máquinas hablando podían cambiar las reglas más rápido.

J: Por ahí podrían haber descubierto redes sintácticas, reglas que nos hubieran parecido piolas, Mirá esto está bueno para incorporar al lenguaje. Pero bueno, parece que los resultados no eran prácticos. Por eso, las IA hoy en día son muy especializadas y lo que yo te decía requiere

mucho del asesoramiento del experto de dominio, depende muchísimo de los datos, dependen muchísimo de las metas y la medición, entonces tal vez necesita expertos para evaluar al sistema, porque vos tenés que evaluar que el sistema esté haciéndolo bien. Entonces vos necesitás un experto que te diga, bueno, vamos a trabajar acá con la asistencia social y la predicción de mujeres embarazadas, y vos tenés que ir viendo con un experto si te está diciendo cosas razonables o el algoritmo se está yendo para el lado de los tomates. Volviendo al principio de la charla y también me trajiste con Kasparov, porque esto es algo que dijo Kasparov que es con el tema del miedo a estas tecnologías. Kasparov a pesar de ser como un emblema del vencido por estas tecnologías, a pesar de que uno puede discutir si era IA o no, en el diario popular Kasparov está como el emblema del gran paso de la IA. Él habla sobre IA, es un entusiasta del tema, él incluso tuvo una compañía que fabricaba máquinas que jugaban al ajedrez por eso digo esto de la Inteligencia bruta, él mismo dice que las computadoras de ajedrez eran fuerza bruta. Él da un concepto muy interesante alternativo, la gente tiene mucho miedo al tema de ser reemplazado, porque uno ve esto, lo ves en el diagnóstico médico, que se empieza a aplicar IA, al diagnóstico de imágenes. Se ha entrenado un sistema con millones de imágenes de lunares, por ejemplo, y clasifica en sospechosos o no, los que dicen bueno, vaya a su casa tranquilo o los que dicen no, vamos a hacerle más estudios, vamos a hacerle una biopsia, o algún estudio complementario y este sistema adquirió un grado de precisión similar a un dermatólogo top. Con lo cual uno se empieza a hacer preguntas sobre el futuro del trabajo, y Kasparov plantea algo muy interesante, que la IA no tiene que ver con un reemplazo sino con una promoción, él dice que la IA es un ascenso para el profesional, porque va a ser como un asistente, lo va a mejorar al médico, no lo va a reemplazar, esa es la visión de Kasparov, en lugar de un reemplazo habla de una promoción de los humanos, que hay tareas más mecánicas, automatizadas que nos las va a mejorar, nos va a bajar el nivel de error y nos va a permitir por ahí concentramos en decisiones más superiores.

F: Sí, ahí el límite va a quedar, porque se entiende que es un asistente si hace fuerza bruta. Si hace fuerza bruta claramente las decisiones finas queda del lado humano, en cuanto el algoritmo toma decisiones...

J: No, por más que sea inteligente como en este caso, está bueno uno como médico tener a alguien que tenga más precisión que vos, o que en caso de duda te diga, realmente me parece bueno. Estaría bueno para mí tener alguien que me diga cuando estoy programando algo si me dijera, no tal enfoque no está.

F: No, igual de todas formas, en el caso de los lunares es los lunares. Solo sirve para lunares, no reemplaza al traumatólogo.

J: Lo interesante es que ahora podemos aplicar la fuerza bruta con aprendizaje. O sea hasta hace poco uno tenía la Fuerza bruta, yo puedo tener acá almacenados 17.000.000 de lunares con sus resultados, si terminaron siendo benignos o malignos pero no hay mucho más por hacer, que es algo superior a lo que el médico más encumbrado puede analizar en su vida, un médico encumbrado puede analizar decenas de miles de casos con toda la furia, y hay que ver si tiene la memoria para acordarse del caso que vio a los 25 años, la computadora tiene una precisión; este lunar de 1977, pasó tal y tal cosa. Y ahora ha sido posible, pero si no veía un lunar que era exactamente igual, no servía, no lo podía aplicar. Ahora en cambio puede definir patrones y puede ver un lunar nuevo y decir bueno, lo más parecido que tengo son estos 17

casos, y de estos 17 casos, 5 hicieron este tratamiento y se murieron, y éstos hicieron este tratamiento y les fue bien, entonces vos decís, ah bueno, vamos por esto, está muy bien por ahí. Por eso yo lo veo como esa idea de Kasparov de que es una especie de promoción, es una ayuda, porque no sólo te puede decir qué tipo de patología es, sino qué pasó con todos los que tuvieron esta patología o muy similares. Ahí ya empezás a cruzar muchos datos, porque vos podés cruzar la edad de cada paciente comparado con el paciente que te vino, no sé, grupo sanguíneo, antecedentes, ahí ya entran un montón de variables que por ahí escapan al cruce humano, entonces bueno, yo tengo 50 variables, edad, peso, antecedentes, diabetes, presión arterial, ves: ah, mirá, acá hay un paciente muy similar que tenía este tratamiento y le fue bien. Por eso digo, están los dos enfoques, los que le tienen miedo a ser desplazados o los que lo ven como un asistente que los va a elevar en su tarea.

F: En esta tención de los dos, antes me decías que en el grupo de IAAR trabajan desde muchos enfoques, psicológico, filosófico, ¿dónde aparecen los cuestionamientos o las tensiones o las discusiones en cuanto a la IA, sobre qué discuten o qué se aparece?

J: Mirá, lo que está más en boga es el tema del futuro del trabajo, como toda revolución esto está siendo y va a ser una revolución industrial, todas las revoluciones en cierto modo funcionaron como promociones para la gente. La invención de la máquina de vapor o la industrialización si bien cortó o eliminó muchos trabajos, después permitió acceder a trabajos de más calidad, a nadie le gustaría volver a los trabajos del siglo 17 que fueron reemplazados por la industrialización, creo... No sé, no estamos con ganas de hacer esos trabajos. La digitalización también, por ahí, no sé, ponerte a pintar puertas de un auto era un trabajo, ya no hay gente, salvo los chapistas que reparan, pero los autos nuevos se pintan con robots, trabajo rutinario, no sé, pegan etiquetas, todos los trabajos que van siendo reemplazados por las cadenas de producción, de montaje. Bueno, acá estamos reemplazando trabajos más interesantes. Los médicos por ahí que el análisis de imagen tengan un asistente por ahí les permite más concentrarse en el tratamiento y otros puntos, no?. Los asistentes sociales tal vez ahora puedan tener sus grupos de trabajo más interesantes para trabajar y poder enfocarse en la detección de causas y tal vez pueden tener mejores resultados puntuales y hacer trabajos más específicos y más jugosos. Lo que yo te decía al avanzar y trabajar información más granulada, Creo que se abre una nueva gama de especialización, yo voy más por el lado de Kasparov, estamos siendo promovidos, entonces hay tareas que vamos a empezar a dejar más en la IA, pero eso también nos va a obligar a nosotros, a partir de ahí, a ser mejores y gestionar esa información. Vamos a ser como gestores o administradores o managers de la IA. La IA de tareas muy puntuales y por ahí nosotros lo que vamos a tener que hacer, mientras no exista la IA general que nos borre, hay que ser como administradores de equipos de trabajos de IA, me imagino así. Después por ejemplo hay conversaciones más humanísticas porque ya se están hablando del futuro del trabajo, en un futuro hipotético que hay que ver, tipo Wall-e, la viste la película Wall-e que las máquinas hacen todo. Claro, entonces a partir de ahí, la idea del sueldo universal tipo la asignación universal, hablamos más de esos temas, o hablan más de estos temas, la parte filosófica, la ética, hasta dónde debería ir la IA. Pero esos puntos a mí me parece que van a quedar como abstractos, la investigación en general no es detenible en general.

F: Si se sabe que va a haber una posible ventaja

J: Sí, pero en general siempre va a haber alguien que va a investigar por el costado, por eso está así la aplicación maligna. También va a estar, así como todas las tecnologías, es inevitable que se utilice o para el hacking, para hacer sistemas inteligentes para violar sistemas de seguridad o para alarmas para el momento, se están haciendo drones inteligentes, misiles inteligentes, sí, desgraciadamente también va a haber eso, las guerras van a tender a ser más robotizadas, uno va a exponer menos a la gente, tal vez sea mejor, porque al final va a terminar siendo un recuento de drones caídos y barcos inteligentes hundidos y bueno, está bien no me quedan más drones, no ya está, ganaste vos y por lo menos te ahorraste la *carnicería*, no sé. Pero sí, en general las industrias de la guerra en general mueven solo en los grandes dinamos de las revoluciones industriales

F: Así entonces tenemos a Elon Musk, insistiendo en la necesidad de regular la IA antes de que sea tarde

J: Sí, están los enfoques, está el enfoque de Zuckerberg que tuvo un gran revés con el tema de información. Claro, él vive del análisis de datos, él vive de la ciencia de datos, el valor de Facebook es la ciencia de datos, nosotros le damos los datos y ellos le ponen el aprendizaje automático encima para que de nuestros datos salgan campañas publicitarias dirigidas. Su negocio, el core de Facebook es hacer ciencia de datos sobre tus datos, o sea el pedido de habeas data a Facebook es decirle "cerrá". Claro, no uses los datos personales: cerrá. Es así, es cerrá o te empiezo a cobrar, o vas a tener que empezar a cobrar. Esa es la función, es el modelo de Facebook. Pero no hay que ser ingenuos, lo sabíamos; hace muchos años que está la célebre frase que dice que "cuando un producto no tiene precio el producto sos vos". Eso lo sabemos.

F: ... Pero ahora se vivió más en carne propia....

J: Debe haber tocado algún interés, nadie medianamente en el tema podía ignorar esto de Facebook. Bueno, hubieron algunas particularidades de violación de privacidad de datos pero hay que ser un poco ingenuos para rasgarse las vestiduras. Pero más allá de eso están estos dos enfoques, de una IA benigna y de una IA potencialmente destructiva. Si vamos a lo que ocurre después de cada revolución industrial, el mundo, a pesar de todo, va avanzando, o sea va mejorando los indicadores de calidad de vida, de guerras, de hambrunas, de muertos por conflictos medidos en proporción al total de la humanidad van disminuyendo, pero esto es una correlación, no una causa, podemos decir que hay una correlación entre la Revolución Industrial y una mejora de la calidad de vida global. Salvo la Argentina, que debe ser el único caso que tenemos más pobreza que hace 50 años y que hace 50, y que hace 100. Pero en general el resto del mundo va como mejorando sus índices, globalmente se ve. Pero aventurar para qué lado puede ir, yo estoy en el medio. Todo esto, sí es aterrador, no por el miedo a lo Terminator, no me imagino todavía tipo los robots toman el control de la tierra, no lo imagino en un futuro de un par de generaciones, no veo. Pero sí por temas de seguridad y sistemas muy inteligentes aplicados para fines no sactos.

F: Se va a usar más para maximizar ganancias

J: Como Facebook por ejemplo.

F: maximizar ganancias, reduzcan personal sin importar las consecuencias

J: Sí, sí, pero ahí yo creo que va a haber... yo confío más en el tema de la promoción, y que la humanidad se va a ir acomodando, como se fue acomodando a otros impactos de pérdida. Yo

creo que en ese sentido, no le tengo tanto miedo al tema de reemplazo, incluso cuando trabajos como el mío puede estar también potencialmente amenazado, porque en algún momento habrá IA que programen, o que generen IA. De hecho hay proyectos prototipos de IA que desarrollan modelos de aprendizaje limitado, pero sobre data, algún planteo, algún norte, unas preguntas, un algoritmo para ambición de eficiencia, ya pueden hacer algo con poca intervención humana.

F: Cuando hablamos de algoritmos y de la IA todo parece que se genera a partir de la idea de dato como lo más relevante

J: No, en realidad todo empieza a partir de la pregunta, yo también cuando doy unas charlas introductorias sobre lo que es IA y aprendizaje automático, cuando uno hace un modelo uno tiene un orden de importancia, de preponderancia que definen el éxito o no de un proyecto, lo que menos define el éxito de un proyecto, en general, es la programación del modelo, la precisión de un modelo, por eso mucho *romperse los cuernos* si el modelo es un poquitito mejor o no, por ahí en algunas cosas puntuales sí, por ahí en detección de imágenes sí es importante, segmentación comercial o esos, modelos bien hechos, va a andar bien.

Pero hay 4 otras fases y son precedentes en importancia. Lo más importante es la pregunta, es saber qué quiero, tener en claro la pregunta, la respuesta que tengo y cómo me dí esa respuesta: yo quiero saber si este lunar es maligno o benigno, sin esa pregunta no sé nada, voy a tener 18 millones de imágenes, y bueno, quiero un modelo que me diga benigno o maligno o que me haga una clasificación de 7 tipos si eventualmente las hubiera. Pero siempre lo importante es la pregunta, ser el experto, por eso siempre el conocimiento del experto y el que está afuera, que traiga preguntas nuevas, para ver si se le puede aplicar IA. Entonces siempre la pregunta. Después, una vez que tenés la pregunta, ahí vienen los datos, sin datos buenos, no existe, el mejor modelo no sirve, si yo te doy cambiada las clasificaciones de los lunares, o en lugar de lunares te mando en el modelo fotos de tacitas de Starbucks, por más que sea el mejor modelo del mundo no sirve. Hay otro dicho dentro de la ciencia de datos que hay una sigla que se llama GIGO, Garbage In, Garbage On: entra basura, sale basura. Entonces primero la pregunta bien formulada y definir qué quiero que me contesten, cuáles son las respuestas. Después está el procesamiento de los datos; el modelo, yo tengo que ver cómo voy a clasificar, qué datos me voy a quedar, si están normalizados, si es texto libre, entonces si es texto libre si yo de alguna forma puedo normalizarlo, llevarlo a valores numéricos para que sea más fácilmente interpretable y clasificable por el algoritmo, detectar datos inválidos, por ejemplo son todos lunares y se colaron 300.000 fotos de vasos, si entran al modelo van a romper todo, porque por ahí ve algo raro y va a decir ah no esto es positivo, entonces detectar eso, limpiar los datos, sacar los datos inválidos, completar los datos incompletos. Por ahí hay un set de datos y falta la clasificación final, ¿qué pasó con esto? ¿fue benigno? ¿fue maligno? ¿qué tratamiento se hizo?. Ahí tenés que tratar de rellenarlo, de ir y averiguar esos casos puntuales ¿qué pasó?, o los sacás. Eso también es importante, porque hace a que entre basura o no al sistema. Esa parte entonces, el preprocesamiento, la selección, los datos, la preparación de los datos, la optimización de los datos es muy importante. Y finalmente sí, está elegir el modelo, de acuerdo al tipo de problema, uno puede optar por distintos tipos de modelos, después dentro de los modelos se puede interpretar de determinado modo, algunos parámetros, se llaman hiperparámetros, evaluarlos, mejorarlos, por

supuesto que es importante, pero es menos importante de los pasos anteriores, es decir un modelo mediocre, pero con una buena pregunta, buen set de datos y un buen preprocesamiento, un buen trabajo de datos, por lo menos te va a dar una buena idea. El mejor modelo, vos traés al mejor, hecho por el jefe del laboratorio de IA de Standoford, le diste mal los datos al capo de deep learning de Standford, o los datos del procesamiento hubo *fruta* ahí adentro, va a hacer agua el modelo. En cambio te lo hace un modelo que está en el Template Azure de Microsoft con buenos datos, va a ser mejor.

F: Para cerrar, hay dos preguntas que te quiero hacer. Una es si viene alguien y te dice te financio tu proyecto soñado...

J: Sí, yo estuve en un par de reuniones con la gente de estas reuniones que hace el Ministerio de Modernización, estuve en un hackaton de la salud, no tanto como un proyecto personal mío, pero yo creo que es un proyecto a donde yo veo una posibilidad fuerte de mejora y de democratización, en la salud. La salud primero que es algo que ya está demostrado, ya se está demostrando que hay avance, no es ciencia ficción. Ya te había comentado, estoy muy insistente con esto de la clasificación de imágenes, pero es un área donde la IA está mostrando avances, donde los privados se están metiendo, al grupo ya se acercaron distintos privados con intereses en desarrollar inteligencia aplicada a la salud, uno era privado y el otro es una fundación de una persona muy famosa. Es decir, los privados están avanzando sobre el tema. En el exterior se está trabajando el tema y creo que podría ser un tema, debería ser un tema de salud pública, que yo creo que con un porcentaje bajo de recursos para lo que es la salud pública, desarrollar sistemas sale una fracción de lo que cuesta por ahí hacer un hospital y además que en el Estado ya hay grupos de gente muy altamente capacitada.

F: ¿Para hacer qué?

J: IA a la salud pública. Elegir diferentes problemas o diagnósticos y empezar a trabajar en IA. Aprovechar, tenemos una gran ventaja, es que tenemos lo máspreciado, que son los datos, tenemos una gran red de salud pública en la Argentina y tenemos que aprovecharlo como pocos países, una de las cosas buenas, por ahí de los que criticamos el tamaño elefantanteásico del Estado es que al tener la salud muy distribuida y pública podemos disponer de un gran set de datos, además, rico con mucha varianza en todas las regiones, prácticamente todos los sectores sociales y económicos. Y eso es muy aprovechable, la verdad que tenerlo ahí sin aprovecharlo da mucha bronca. Y sobre todo que va a permitir la democratización, porque la ramita de IA puede ayudar a llevar una salud de alta calidad y estandarizada a cualquiera. Nosotros tenemos IA para hacer diagnóstico y que pueda ser consultada por cualquiera que tenga acceso a internet, por cualquier médico que tenga acceso a internet, que pueda subir una imagen y es experto en información en IA, pueda dar un diagnóstico orientativo, le hacés llegar salud especializada, de calidad a cualquier punto del país y a personas de cualquier nivel económico. Porque también tenemos que los lugares alejados uno no tiene acceso a los especialistas, o tiene que sacar un turno en una ciudad a 200 km para dentro de un mes, por ahí uno tiene un médico clínico, pero uno no tiene un neumonólogo, entonces un médico clínico ve una imagen de un pulmón, y salvo algo muy evidente no sabe qué hacer.

F: Hay algo muy similar al poder democratizador que tendría esto que es la educación en línea que permite llevar una educación de calidad a cualquier punto sin depender de ...

J: Exacto, vos podés tener un diagnóstico top en lo más recóndito del país, el ahorro de tiempo y de recursos, porque el médico clínico ya con eso te puede hacer las órdenes o te puede agilizar el trámite, te puede decir se sugiere tal y tal estudio. Y vos después vas a ir a ver el experto a 200 km pero a lo mejor ya te pudiste hacer los estudios y es más rica esa consulta, ya no es que lo vas a ver al experto y te va a decir “ah ahora hacé tal y tal estudio y vuelva a verme dentro de un mes”.

F: La pregunta es, el diagnóstico por imagen, ¿está desarrollado, se compra ya como algoritmo?

J: Primero que sí, pero bueno, aquí está el tema, el tema es que tenemos gente para desarrollarla y no anda pagándole un royalty a otra persona. Porque ¿qué va a pasar? va a haber gente que lo va a desarrollar y la salud pública lo va a usar en algún momento, cuando esto sea una herramienta de uso común, como fue en algún momento el tomógrafo, la resonancia magnética, primero lo tendría no sé, el Hospital Austral supongamos o no sé la Clínica del sol hablando de los '80, y después fue llegando a los hospitales públicos, la diferencia es que alguien lo va a hacer, y alguien lo va a cobrar, y no creo que lo cobre barato.

F: Ahí está la misma tensión de software libre, software propietario, uno tiene que apuntar a que haya modelos abiertos...

J: Exacto y ¿cuál es el punto acá?. La materia prima para estos sistemas son los datos, y acá tenemos que el Estado argentino es dueño de esos datos, o sea que tiene la materia prima. Un temor mío es que venga algún privado, compre la exclusividad de uso de esos datos, desarrolle un sistema y después te venda a vos el uso del sistema. Es que incluso ni siquiera puedas usar esos datos porque firmaste un acuerdo de exclusividad con esos datos, y yo creo que es algo además desarrollable, con un grupo determinado de personas que ya está en el Estado, ya han contratado científicos de datos a distintas áreas, por ahí con el apoyo de algún privado, me da la impresión que lo veo un proyecto realizable, es decir, está la materia prima, está la mano de obra, no me da la impresión que sea un presupuesto, de hecho sería por ahí hasta reutilizar gente que ya le estás pagando el sueldo, ni siquiera estoy hablando de hacer una erogación adicional si no simplemente un reacondicionamiento de recursos. Se puede ir avanzando en esto. Esto creo que sí es un futuro, y lo que te decía, creo que va a permitir, democratizar y hacer llegar medicina de calidad a cualquier rincón. Eso yo lo veo como algo hacia donde debería ir ya, a mí me parece una gran oportunidad y que no terminemos después pagándole un privado por eso, o a alguien haciendo un negocio con eso, porque yo creo que la salud lo va a terminar usando, y la salud pública en algún momento lo va a terminar usando. Entonces tenemos la opción de *hacer punta* nosotros.

F: considerando que se trata del resguardo, de un Estado que protege los datos de los ciudadanos...

J: Sí, por ahí lo veo más mercantilista, yo lo veo como una oportunidad de hacer algo y ahorrar más plata, quiero decir, que nos llegue antes, y nos ahorre plata, que no haga un privado negocio con esto, y que además incluso podría ser un activo valioso para el Estado que desarrolla un sistema con una base de datos pública tan rica, hasta incluso podría después venderlo o cobrar el servicio a otro Estado o a privados. Yo lo veo como una gran oportunidad, tanto desde mejorar la salud hasta de hacer un negocio por parte del Estado.

Otra gran oportunidad que veo que otros países desarrollan que acá no tanto es el tema la IA

aplicada al agro. Esto lo veo más para los privados que también tuvimos consultas en el grupo, hay distintas aplicaciones o potenciales usos, por ejemplo análisis de imagen, para detectar el estado de cultivo, ahora que se están utilizando mucho los drones que sacan fotos geolocalizadas, con el análisis de imágenes, comparación con un historial de imágenes detectar, si tienen los cultivos algún tipo de enfermedad, deficiencia, volvemos a lo mismo, clasificación, un diagnóstico, y decirle bueno acá corresponde aplicar herbicidas, sí, no, qué cantidad, más riego, menos riego. Ese tema de monitoreo de cultivo por análisis de imágenes con IA, es algo que se está empezando a usar en el mundo y acá estamos medio *en veremos*, pero te vienen con la idea. El agro también es un ámbito que nos está consultando y donde también potencialmente se puede hacer, porque en otros lados también está por ejemplo el tema de los tractores autónomos, acá todavía no. Ahí está, se manejan tractores, y sí, en ese sentido. Además un tractor no es lo mismo que un auto autónomo en la ciudad. Tenés un campo, da vueltas, vas, venís, vas, venís, y tendrás que tener algún radar para detectar si se te cruzan objetos, frenar, no mucho más que eso. Pero ese es otro terreno donde va mucho.

F: Para terminar te pido que nos cuentes las series que recomendás. Lo hemos hecho con los otros entrevistados, como marca, series que hayas visto, que te hayan interesado, que te llamaron la atención.

J: Bueno no voy a ser original, bueno, Black Mirror por supuesto, no voy a ser muy original, pero miro The Walking Dead con mis hijos, Expedientes X también. Hemos visto La casa de papel, Merlí, somos muy de Netflix, por supuesto las clásicas, Breaking Bad, House of Cards, Cold me gustó mucho, Homeland, Mad Men, saliendo un poco, es clásica, ya tiene un tiempo, pero es una joya, como está la concepción que tiene. Vi una temporada de una que era danesa, era una símil Merlí pero de una profesora...

F: Rita

J: Si, Rita, está muy buena

F: Bueno, gracias

J: No, por favor