

Machine Learning Operations

DOI: 10.29236/sistemas.n165a5

Los desafíos de su implementación en Colombia.

Sara Gallardo M.

La producción de datos es cada vez mayor y su aumento seguirá en auge en computadores, teléfonos inteligentes, tabletas y la variedad de equipos que circulan entre los diferentes usuarios de la tecnología informática y de las comunicaciones, de ahí la importancia de profundizar en el tema.

Para tratar los asuntos más relevantes de este entorno en el país fueron convocados varios profesionales, quienes a través de su experiencia se ocuparon de analizar diferentes aspectos relacionados

con “Machine Learning” y su avance en el país.

El editor técnico y el coeditor de esta edición abrieron el debate.

Joseph Alejandro Gallego M.

Editor Técnico

Revista Sistemas Edición 165

Fabio Augusto González O.

Coeditor Técnico

Revista Sistemas Edición 165

En DevOps llamamos muro de la confusión al que se crea típicamente entre los desarrolladores de

software y los ingenieros de soporte del software ¿En su experiencia, en qué consiste el muro de la confusión en proyectos de machine learning y cómo MLOps ayuda a tumbar ese muro?

Raúl Ramos P.

*Profesor e investigador
Universidad de Antioquia*

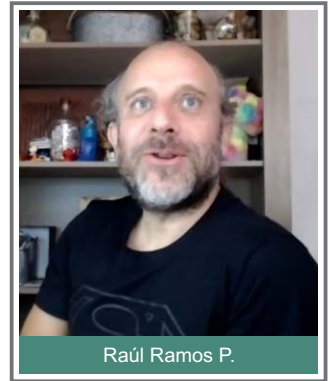
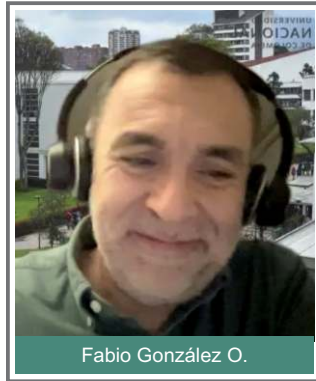
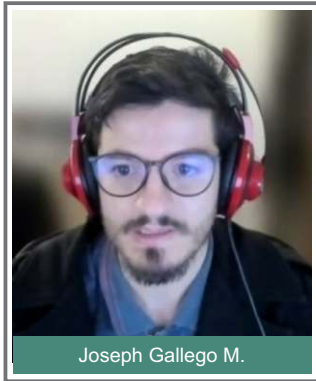
Con frecuencia hago una analogía que en esta ocasión me parece muy pertinente. Hoy todos estamos aprendiendo sobre cómo usar en la práctica y de manera útil todas estas técnicas de IA. El DevOps –sorprende que sea tan reciente- estaba en proceso de implementación

en ingeniería civil, mecánica, en una red de análisis y en prueba en los primeros proyectos de software. Hace cuarenta años había muchos problemas porque los requerimientos no eran del todo claros y se desarrollaron las metodologías ágiles. Con Machine Learning e IA está sucediendo lo mismo.

Ricardo Barragán O.

*Consultor Sénior
Teradata*

En el mundo de la analítica hay dos áreas que trabajan un poco en contravía; normalmente son áreas con personas innovadoras que están a la vanguardia, explorando



constantemente; no están basadas en procesos, sino en resultados. De ahí que un científico de datos no siga un proceso sino busca un resultado, sin importar qué herramientas tenga para lograrlo. Y, por el otro lado, están las áreas de tecnología, las cuales tienen que propender por la seguridad de una compañía, por la información, los ciclos, los tiempos y están basadas en el proceso. Se trata de dos áreas que trabajan en sentidos distintos y MLOps busca que trabajen juntas. Es decir, que exista la innovación, la flexibilidad, que la operación sea algo organizado y basado en procesos, es una forma de ver esas barreras.

Y otra forma que es importante mencionar, lo cual me recuerda un paper de Google que muestra cómo desarrollar un modelo analítico, es una parte pequeña con respecto a todo el proceso que se requiere para operacionalizar un modelo; es decir, no necesariamente cuando se termina de desarrollar un modelo, este es usable para la compañía.

Fabio A. González O.

Me parece muy interesante la perspectiva de Ricardo, en esa época ¿cómo se manejaba esa tensión, existía? Y ¿cómo se resolvía en proyectos de inteligencia de negocios, de ingeniería de datos?

Ricardo Barragán O.

Esa parte ha cambiado un poco; antiguamente, el 99% de los mo-

delos del área comercial, eran modelos descriptivos, principalmente era una visualización en una gráfica o en un reporte; hoy son otros y se trata de modelos predictivos o prescriptivos. Eso cambia un poquito la perspectiva y la operacionalización del modelo; se requiere tener en cuenta el aprendizaje y la ejecución del modelo. En mi compañía lo llamamos “scoring”, no se cómo lo denominen en la industria.

Fabio A. González O.



Es decir ¿eso tiene que ver con la complejidad de los modelos? ¿Esa brecha se ha agrandado?

Ricardo Barragán O.

Desde mi punto de vista sí hay más opciones, más herramientas para tener analítica y ahora las más usadas y las más difíciles de operacionalizar son las que contienen inteligencia artificial, el aprendizaje del modelo.

Diego Ibagón P.
*Machine Learning Senior Manager
Mercado Libre*



En compañías muy grandes de diferentes sectores el miedo inherente a llevar esta producción sigue existiendo y es curioso. Se explica también por la idiosincrasia. La relación academia empresa en los grupos empresariales grandes es muy lenta y nos quita la velocidad que sí tienen otros países, entonces ese miedo inherente genera el uso de enfoques antiguos, hace que uno trate de meter un triángulo en los juegos de los niños o un triángulo en la caja del cuadrado o el círculo en otra que no cuadra.

Jorge E. Camargo M.

El muro de la confusión en el ML-Ops es todavía mucho más complejo, es decir, estamos en una fase en que las empresas empiezan a adoptar esto. Si bien ya se está entrando a una adolescencia en el te-

ma del software, porque todavía no se puede decir adultez, se han organizado de alguna manera la casa y los roles; el tema todavía no se comprende muy bien, se especula mucho. De manera que un muro compuesto por ingeniería de software, más ciencia de datos, analítica de datos es pedregoso. Se trata de un tema en el que todavía estamos en pañales y en las entidades no se entiende muy bien.

Raúl Ramos P.



Me parece muy pertinente lo relacionado con los equipos multidisciplinares y la parte del miedo, del pánico a la producción y creo que no es solamente una cuestión de cómo nivelar los equipos que están compuestos de muchas profesiones distintas; también es una cuestión cultural de cada profesión; no es igual, la parte del conocimiento, la manera de trabajar de un ingeniero de sistemas frente a la de un

físico o un economista. Cada vez que en un proyecto no está definido lo que hay que hacer, es necesario interrelacionar con el cliente y una variedad de profesiones. Es más viable en lugar de ingenieros de sistemas, optar por vincular físicos, en la medida en que los primeros requieren los asuntos muy bien definidos. No saben asumir el vocabulario de los otros ni manejar una terminología distinta a la propia. Esta cuestión de los equipos multidisciplinarios es muy difícil de poner en marcha.

Joseph Alejandro Gallego M.



En ingeniería de software se utiliza el término deuda técnica para referirnos a las consecuencias que conlleva el reprocesso, debido a soluciones rápidas a problemas complejos. Esto se debe a presiones internas y externas al equipo de trabajo.

En su experiencia como experto ¿cuáles son los principales riesgos

en los proyectos de Machine Learning que desencadenan en deuda técnica?

Ricardo Barragán O.



En mi opinión, veo los problemas de inteligencia artificial y Machine Learning un poco diferentes a los problemas de desarrollo que teníamos antes; de hecho, son más restantes. En todas las organizaciones se ve que el científico de datos está trabajando con una versión y el otro está trabajando con una diferente y el problema lo van a tener todos en el momento en el que quieran operacionalizar, a menos de que se haga independiente esta acción de los desarrollos. Debe haber una capa en el medio que independice eso.

Raúl Ramos P.

En mi opinión el concepto de deuda técnica aplicado a este tipo de entornos es un poco distinto también;

en la línea de lo que se está mencionando es bueno hacer ciclos rápidos, equivocarse pronto, gestionar las expectativas de los clientes, salir a producción lo antes posible, aunque sea parcialmente. El riesgo es no darse la oportunidad de hacer todo eso, porque en muchas ocasiones esto es también un ejercicio de descubrimiento, de validación de cuáles tecnologías y métodos pueden resolver el problema.

Fabio A. González O.

Muchos científicos de datos desarrollan buena parte de su trabajo usando herramientas interactivas como los cuadernos de Jupyter. ¿Cómo pueden estos profesionales evolucionar su trabajo hacia el uso de MLOps?, ¿en qué aspectos se deberían enfocar primero?

Jorge E. Camargo M.



Lo que uno encuentra con estos procesos tan nuevos en las com-

pañías es que los roles todavía no son claros, no están definidos y, muchas veces, las primeras personas que empiezan a abordar estos proyectos hacen de todo, es decir resultan aprendiendo machine learning, conectándose a la nube, subiendo máquinas. En otras palabras, haciendo de todo, porque no hay más profesionales en las compañías.

Raúl Ramos P.

En mi opinión, se trata de dos caminos muy distintos en términos de formación. Me parece que los profesionales deben acercarse para conocerse entre sí, es necesario interactuar para interpretar los datos y los modelos.

Ricardo Barragán O.

La falta de conocimiento profundo de las ramas requeridas siempre va a repercutir en costos. En una oportunidad trabajamos para un cliente con un científico de datos, una eminencia con amplios conocimientos, pero el consumo de los recursos era enorme y resultaba muy costoso para la compañía, de manera que se optó por un ingeniero que depuró el programa y quedó consumiendo el 80% menos de recursos. Es necesario el trabajo conjunto, cada uno con su aporte individual.

Diego Ibagón P.

Es importante generar conciencia en términos de costos y para ello es aconsejable trabajar con ejercicios muy didácticos de cara al diseño.

Fabio A. González O.

Sabemos que se trata de un tema nuevo del que todos estamos aprendiendo. ¿Qué se necesitaría para que una empresa pueda empezar a cambiar esa cultura? Desde las universidades a nivel de Ingeniería de Sistemas y otras áreas afines ¿qué nos faltaría?

Diego Ibagón P.

El asunto no es fácil, no se me ocurre la solución completa. Los ingenieros de sistemas adolecemos de muchas cosas. Los jóvenes, además de ver la solución en una forma más práctica, tienen conciencia del impacto y en ese sentido hay que volver a lo básico en muchos aspectos.

Raúl Ramos P.

Desde la perspectiva de la formación, en la Ingeniería de Sistemas es necesario involucrar la cultura de “poner cajitas juntas”, además de profundizar en diferentes aspectos. Así mismo, dar lugar a la prueba y el error. En las empresas es

importante entender que se trata de proyectos de largo aliento y que los resultados pueden demorar.

Ricardo Barragán O.

No estamos viendo los problemas a largo plazo sino simplemente el día a día en cuanto a la formación. En cuanto a las empresas deben innovar, aprender porque el mercado es quien valida el producto desarrollado.

Jorge E. Camargo M.

En todas las profesiones tenemos un déficit de formación importante, en muy pocos colegios se desarrolla el pensamiento computacional, asunto clave sin importar la carrera. En esa dirección, los retos son grandes.

Ricardo Barragán O.

Desde mi experiencia como proveedor, en primera instancia se debe analizar el problema y luego elegir la tecnología para solucionarlo. Por lo general, esto no sucede y el camino se recorre al revés. 🌐

Sara Gallardo M. Periodista comunicadora, universidad Jorge Tadeo Lozano. Ha sido directora de las revistas Uno y Cero, Gestión empresarial y Acuc Noticias. Editora de Aló Computadores del diario El Tiempo. Redactora en las revistas Cambio 16, Cambio y Clase Empresarial. Coautora del libro “Lo que cuesta el abuso del poder”. Ha sido corresponsal de la revista Infochannel de México; de los diarios La Prensa de Panamá y La Prensa Gráfica de El Salvador y corresponsal de la revista IN de Lanchile e investigadora en publicaciones culturales. Se ha desempeñado también como gerente de Comunicaciones y Servicio al Comensal en Inmaculada Guadalupe y amigos en Cía. S.A. (Andrés Carne de Res) y editora de Alfaomega Colombiana S.A.; en la actualidad es asesora en escritura y producción de libros. Es editora de esta revista.