mdt.

Ex Datis, Claritas

El gran "mapa" del Machine Learning, la IA y los Datos, edición 2024

Como cada año desde hace diez, Matt Turck publica el mapa que clasifica a los principales jugadores del mercado de datos, machine learning e IA. Edición 2024.

Manuel Delgado Tenorio · 7 de abril de 2024

Publicado originalmente en www.manueldelgado.com/blog/el-gran-mapa-del-machine-learning



Como cada año desde hace ya diez, Matt Turck ha publicado el informe en el que recoge y clasifica a los jugadores más importantes del mercado de la tecnología de datos.

Con más de 2.000 logotipos, el mapa resulta impresionante, sin duda. Ha crecido en casi 500 empresas en solo un año, un crecimiento impulsado por olas masivas de creación de compañías y financiación, especialmente en infraestructura de datos y ML/AI.

Sin embargo, como suele ocurrir todos los años, **lo más relevante del "_landscape_" de Matt Turck no es el mapa.** Si de verdad queremos entender qué está pasando el mercado de la IA y los datos, la verdadera sustancia está en el texto que acompaña al gráfico. Menos sexy, pero mucho más útil.

¡Gracias por leer It's All About the Data! Suscríbete gratis para recibir nuevas publicaciones y apoyar mi trabajo.

De los numerosos comentarios incluidos, yo me quedo con la discusión sobre la evolución del concepto de *Modern Data Stack* (que podríamos traducir como "Plataforma de datos moderna") hacia el concepto de *Modern AI Stack* ("Plataforma de IA moderna"), más cercano a la realidad actual, al menos a primera vista.

Como bien indica Turck en su informe, el salvaje aumento de interés hacia la IA generativa que estamos viendo desde hace año y medio ha permitido el nacimiento o el crecimiento de un gran número de compañías que resuelven el problema de cómo alimentar con datos a las numerosas aplicaciones de IA generativa que se están desarrollando. De ahí que se empiece a hablar de *Modern AI Stack* como la evolución de la plataforma de datos "convencional", más orientada al dato estructurado.

Personalmente, soy incluso más cauto que Turck al abordar este tema. El principal motivo para mi cautela es que ese aumento en el número de compañías que ofrecen esta solución no guarda relación, en este momento, con el valor de mercado de resolver el problema, sino que responde más al *hype* alrededor de la IA generativa y a las expectativas de la industria (en mi opinión, muy infladas) sobre el tamaño futuro de ese mercado.

En la práctica, hoy, el porcentaje de empresas de tamaño representativo que han llegado a **poner en producción** aplicaciones de inteligencia artificial generativa es muy bajo, entre el 10% y el 25% dependiendo de qué analista te dé la cifra. Y, de esas aplicaciones, las que usan datos no estructurados propios de la organización con actualizaciones frecuentes son, en mi experiencia, un porcentaje mínimo.

Así que la inmensa mayoría de las empresas siguen operando bajo el paradigma de la "plataforma moderna de datos" sin que nada les obligue en el corto o medio plazo a evolucionar sus capacidades de gestión a gran escala de datos no estructurados.

Con muchísima probabilidad, para cuando llegue el momento en que realmente esa necesidad esté generalizada, los grandes jugadores del mercado de plataformas de datos modernas habrán incorporado esas capacidades que hoy abanderan las nuevas empresas de nicho y, casi sin darnos cuenta, comenzaremos a considerarlas una funcionalidad más de cualquier plataforma de datos.

Esto ya está ocurriendo, por ejemplo, con la posibilidad de almacenar vectores en bases de datos: algo que hace un año apenas podían hacer unas pocas startups se ha incorporado ya como nueva funcionalidad en muchos de los motores de bases de datos más ampliamente usados.

Veremos lo mismo con prácticamente todas las capacidades que hoy parecen pertenecer a esa nueva categoría de la “plataforma de IA moderna” y, en consecuencia, la categoría se diluirá tanto que desaparecerá.

Si quieres leer el informe completo y ver el mapa en un formato interactivo muy práctico, puedes hacerlo en la web de Matt Turck.

P. D. - Es genial ver cómo empresas españolas como Graphext aparecen en el mapa de Matt Turck. Un gran éxito de Victoriano Izquierdo y de todo su equipo. Si no conoces Graphext, echa un vistazo a su web.



Sobre el autor

Manuel Delgado Tenorio

Manuel Delgado Tenorio es asesor independiente de estrategia de datos e inteligencia artificial, profesor e investigador. Cofundó y vendió una *startup* de inteligencia artificial, fue socio de la práctica de Data & Analytics para servicios financieros en EY y responsable de Data Analytics para el sur de Europa en AWS. Es profesor asociado en IE University, donde imparte la primera asignatura de grado sobre Customer Lifetime Value en España, investiga sobre el CLV en su doctorado en la Universidad Complutense de Madrid y es autor del libro "Customer Lifetime Value (CLV): No todos tus clientes valen lo mismo".

manueldelgado.com · LinkedIn · X

Recibe cada artículo en tu correo con la newsletter Ex Datis, Claritas: manueldelgado.com/newsletter/suscribete