

Primeros pasos con IA

Cómo Lenovo e Intel impulsan aplicaciones prácticas entregando IA más inteligente para todos.



Lenovo ThinkSystem SC750 V4 Neptune®, con procesador Intel® Xeon® 6 - Desempeño y eficiencia sin compromisos.

Smarter
technology
for all

Lenovo



Tabla de contenidos

- 3 El rápido crecimiento de la IA
- 4 Habilitando la IA en todos los lugares
- 5 Lo Básico
- 6 Inferencia de IA
- 7 Desbloqueando Insights
- 8 Expertise en transformación de IA
- 10 Acelerando la implementación
- 11 Estudio de Caso: Experiencias del Espectador
- 12 Flexibilidad para escalar sin problemas
- 13 Olhando para a sustentabilidade
- 14 Una Abordagem más Inteligente





El rápido crecimiento de la IA

La Inteligencia Artificial (IA) ha dado pasos tremendos desde sus días pioneros en la década de 1950. Los algoritmos estáticos predefinidos diseñados para análisis estadístico y predicción, ejecutados en los primeros computadores, dieron lugar a los primeros ejemplos de aprendizaje de máquina en la década de 1980, cuando los algoritmos fueron enseñados a reconocer relacionamientos y construir modelos de sistemas complejos.

El advenimiento de grandes redes neuronales en los años 2000 abrió camino para expansiones masivas de capacidad computacional y la introducción de IA generativa y grandes modelos de lenguaje capaces de trabajar con patrones complejos y abstractos.

Desde una perspectiva de negocios, el potencial para derivar insights, reducir cargas de trabajo y acelerar la productividad parece casi ilimitado y las empresas están investigando ativamente maneras de poner la IA para funcionar...

89%



de los líderes de TI dicen que están pesquizando o usando tecnología habilitada para IA¹.





Habilitando la IA en todos los lugares

Lidar con iniciativas de IA puede ser asustador. Históricamente, la IA ha estado apenas en el reino de los motores de búsqueda, instituciones financieras y pesquisa científica. Além del costo de adquisición del hardware de computación en sí, en muchos casos, los data centers existentes no consiguen suportar los requisitos adicionales de energía y resfriamiento, lo que exige tiempo y despesas de capital adicionales.

La buena noticia es que la introducción de modelos de IA de base amplia, treinados en datos públicos, disminuiu las barreras para las organizaciones implementarem soluciones avanzadas de IA.

Lenovo e Intel están poniendo su parceria de longa data para funcionar, fornecendo soluciones que permiten que las empresas aprovechen todo el ótimo trabajo que fue hecho hasta ahora y apliquen la IA de maneras prácticas que entreguen resultados mensurables.

Vamos empezar con lo básico

En términos más simples, la IA es largamente definida como cualquier sistema de automatización que simule la inteligencia humana aprendiendo en el trabajo. La IA es implementada en dos fases:

1

Treinamiento o desarrollo de modelo

Este es el proceso en que los científicos de datos desmenuan y optimizan modelos básicos con un conjunto de datos seleccionado...

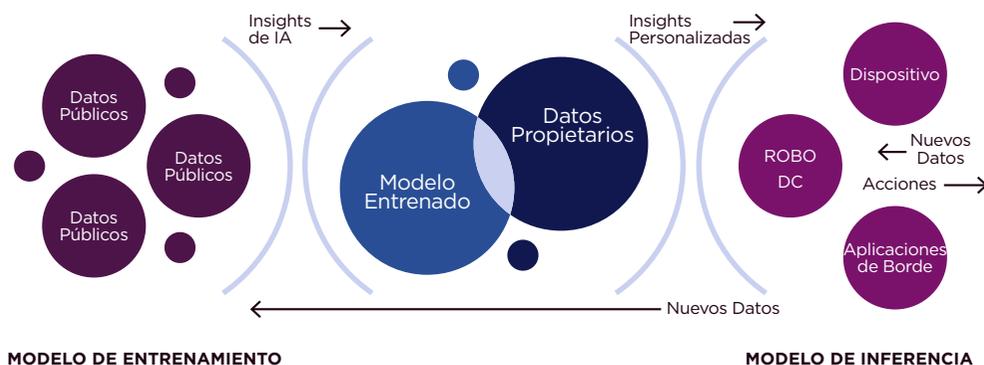
2

Inferencia

Aplicando nuevos datos a un modelo treinado para obtener nuevos insights y acelerar la automatización.



Sistemas de Acción



Desenvolvimiento de modelo de entrenamiento

El entrenamiento es alcanzado por medio de un proceso llamado aprendizaje de máquina (ML), donde un modelo es treinado con base en parámetros específicos que definen la tarea (por ejemplo, color, formas y bordas) y usa técnicas como clusterización, regresión y red neuronal que procesan enormes cantidades de datos para desmenuar previsiones. A partir de ahí, el modelo continua a consumir y analizar datos en cuanto melhora su comprensión de esos recursos.

Los conjuntos de datos usados para el treinamento de modelos básicos crecieran a una escala que requiere grandes cantidades de hardware de computación especializado que depende de millares de procesadores ejecutados en paralelo para entregar los recursos necesarios. Es por eso que el treinamento de modelos básicos de IA ha sido tradicionalmente el dominio exclusivo de pesquisa académica, financiera y governmental



La cantidad de poder de computación necesaria para treinar los mayores modelos de IA está doblando la cada

3 a 10
meses.²



Apresentando la inferencia de IA

La inferencia de IA involucra pegar modelos entrenados existentes y aplicá-los a nuevos conjuntos de datos propietarios para tareas específicas de aplicación. Los outcomes y los insights son entonces adaptados para nuevas aplicaciones que son personalizadas para entregar experiencias más precisas y relevantes.

La inferencia se basa en el aprendizaje ya alcanzado, entonces las demandas de procesamiento de generar previsiones e insights son significativamente menores del que aquellas requeridas durante el entrenamiento inicial.

Com los requisitos de procesamiento disminuídos, la inferencia de IA está abriendo las puertas para que más negocios y organizaciones de todos los tamaños aprovechen el poder de la IA para una larga gama de aplicaciones.



Veja como a Lenovo e a Intel estão acelerando a Indústria 4.0 com inspeções visuais assistidas por IA e análises de dados mais rápidas.

Saiba mais.



El número de empresas que usan IA crece

300% en **5 años.**³

Como las organizaciones no precisan desentender los modelos de entrenamiento básicos, eso acelera el desarrollo drásticamente y los ayuda a moverse hacia aplicaciones reales más rápidamente. Adicionalmente, como la abordagem precisa apenas de nuevas informaciones aplicadas al modelo, los datos y demandas de procesamiento pueden extenderse además de los *data centers*. Eso significa que la inferencia puede acontecer donde los datos son colectados, incluyendo en la *borda*.

Eso es importante porque permite el control en tiempo real de funciones críticas que no incurrir en penalidades de latencia yendo y viniendo para la nube, como los sistemas autónomos encontrados en carros autónomos o fábricas automatizadas.

El modelo de inferencia de IA entrenado funciona apenas con los datos necesarios para tomar las decisiones, lo que acelera el proceso de decisión y reduce la necesidad de mover grandes cantidades de datos entre las redes.

Por ejemplo, en un ámbito de manufactura, los servidores de *borda* en la línea que ejecutan modelos de inferencia de IA pueden usar visión computacional (basada en modelos entrenados existentes) para identificar defectos, tomar decisiones y tomar la acción apropiada (usando datos locales propietarios) para tratar con el defecto, manteniendo la productividad de la línea.

Desbloquee *insights* en seus datos más rápidamente

En cuanto las aplicaciones de IA mejoran, la tecnología que las suporta está mejorando para se adaptar a esas nuevas expectativas, en cuanto acelera e capacita la implementación de IA en cada etapa, en la borda a la nube. Lenovo e Intel se unieron para entregar soluciones construídas para fines específicos, proyectadas específicamente para aplicaciones de inferencia de IA.



La última generación de servidores ThinkSystem, como el **Lenovo ThinkSystem SC750 V4 Neptune**, alimentado por procesadores Intel® Xeon® 6, proyectados para IA. Intel® Xeon® ofrece el desempeño con eficiencia energética y recursos de sustentabilidad para ayudar a desminuir el costo total de propiedad. La aceleración integrada entrega un desempeño aumentado para tareas de inferencia de IA y reduce los requisitos de energía y resfriamiento, lo que significa que los servidores Lenovo ThinkSystem SC750 V4 Neptune pueden ser implementados en *data centers* existentes em vez de construir nuevos centros.



5.5x

mejor desempeño de inferencia de IA, del que procesadores concurrentes⁴.



100%

de remoción de calor del resfriamiento líquido Lenovo Neptune de 6ª Geração⁵.



40%

reducción en el consumo de energía.

Con la IA en la borda, las organizaciones pueden capitalizar informaciones dinámicas en tiempo real y entregar mayor automatión, remediación e *insights* donde es más accionable - en la línea de frente.

Aprovechando la *expertise* en Transformación de IA

Proyectar e implementar modelos de inferencia de IA que entreguen *insights* confiables y accionables exige un conjunto muy específico de habilidades y extrema atención a los detalles.

El **Lenovo AI Discover Center of Excellence** reúne especialistas en IA de Lenovo e Intel para ayudar sus desenvolvedores a crear y acelerar la entrega de aplicaciones de IA y modelos de inferencia de IA.



Nuestros especialistas en IA conducen una grande gama de workshops para fornecer avaliações abrangentes de negócios, avaliações de TI e planos de design documentados.



Engenheiros técnicos, socios y cientistas de datos optimizan sus códigos de IA usando *frameworks* de código abierto para ejecutar en los servidores ThinkSystem con *hardware* y *software* Intel.



Podemos ayudá-lo a aprovechar el conjunto abrangente de recursos de la Intel como el toolkit OpenVINO™ y la oneAPI Deep Neural Network Library (oneDNN) para simplificar la implementación de la inferencia de aprendizaje profundo para centenas de modelos pré-treinados.

Lenovo también ofrece una grande gama de workshops de Servicios Profesionales Lenovo para acelerar su jornada de transformación de IA.



El Toolkit OpenVINO™

Las barreras a la adopción de IA generalmente incluyen la necesidad de modelos grandes, optimizados y diversos, una amplia gama de arquitecturas xPU (muchas veces implementadas juntas) y un ecosistema expansivo de *frameworks* de *software* de API para elegir. La implementación de IA puede ser un proceso difícil y demorado, envolviendo muchas elecciones de ecosistema de proveedores.

Con toda esa complejidad, las pruebas de concepto muchas veces nunca llegan a la producción, creando un “cementerio de POC”.

Esas barreras precisan ser derrubadas para crear oportunidad, y es eso que el **OpenVINO™** faz, ofreciendo un toolkit de código abierto que soporta una amplia gama de arquitecturas xPU y *frameworks* de *software* de IA.



Beneficios:

1. Amplia accesibilidad para múltiples arquitecturas xPU a través de un modelo de código abierto.
2. Una solución de inferencia de IA accesible y eficiente que reduce los costos de adopción y aplicación de la tecnología de IA de la borda a la nube para PCs locales.
3. Una arquitectura abierta que permite colaboración en todo el ecosistema—desde científicos de datos criando modelos para *frameworks* de aprendizaje profundo hasta desenvolvedores de aplicaciones en una variedad de verticales, utilizando funciones de IA multimodal de visión, procesamiento de lenguaje natural, sistemas de recomendación e IA generativa.

Emparejado con la **Intel® Edge Platform**, soluciones completas nativas de la borda pueden ser construídas para acelerar iniciativas de IA de la borda con recursos de entrenamiento, optimización y desenvolvimiento de aplicaciones para modelos de IA.

Las empresas también pueden integrar y gestionar con seguridad una flota de nodos de borda, aprovechando los componentes brownfield o greenfield más adecuados y económicos en parceria con nuestro ecosistema incomparable para menor costo total de propiedad.



Vea cómo Lenovo e Intel están facilitando la adopción de la IA con OpenVINO™

Más información.

Acelera tu jornada con soluciones de implementación comprobadas

Cuando llega a la hora de implementar su solución de inferencia de IA, el programa Lenovo AI Innovators agiliza el proceso con soluciones comprobadas usando el mejor software ISV de la categoría en la infraestructura optimizada para IA de Lenovo y de la Intel.

Lenovo e Intel construyen, testan y validan soluciones de inferencia de IA con un ecosistema de socios de AI Innovators comprobados para garantizar implementaciones suaves e ideales que los mantienen en el cronograma y en el presupuesto.

- ✓ Solución de gerenciamiento remoto de **Nybl**.
- ✓ Solución de inspección visual auxiliada por IA de **byteLAKE**.
- ✓ Soluciones de visión computacional, manutención predictiva y detección de anomalías de **Guise AI**.
- ✓ Solución de Análisis de Filas y Multidão del **WaitTime**.
- ✓ Solución de la **Sunlight.io** que acelera la transformación digital de restaurantes y *drive-thrus*.
- ✓ Solución de inteligencia industrial **Smartia** que conecta y transforma datos en percepciones accionables.

Continuamos a monitorar, avaliar y construir relacionamentos con socios ISV conforme sus soluciones evolucionan.

Estudo de caso: La IA está transformando las experiencias del espectador

Lenovo y **WaitTime** revelan una solución innovadora para locales para transformar la experiencia del espectador de la Fórmula 1® usando tecnología de punta. Ao Combinar 18 cámaras estratégicamente instaladas en todo el autódromo del Circuit of The Americas (COTA) con la tecnología de IA patentada de la WaitTime em servidores Lenovo ThinkEdge alimentados por processadores escalables Intel® Xeon®, los operadores de la COTA pueden monitorar meticulosamente grupos de personas en filas.

“Esta plataforma de análisis de datos en tiempo real forne insights inestimables, permitiendo que los operadores entiendan dinamicamente como las multidoes están creciendo, se moviendo y mudando”, disse Zachary Klima, fundador y CEO de la WaitTime.

“Tales informaciones instantáneas capacitan-nos para fazer ajustes inmediatos en las operaciones e estrategias de receita, garantindo una experiencia ideal y continua para los espectadores, en el mismo momento en que maximizan la eficiencia y la receita del evento.”



Puedes ler más sobre la solución **clikando aqui.**



Gana la flexibilidad para escalar perfectamente

La implementación de la inferencia de IA requiere mucho menos en despesas iniciales en comparación con la construcción y treinamento de modelos básicos desde el cero, mas aún existen custos a ser considerados para *hardware, software* y servicios.



Lenovo TruScale ofrece la flexibilidad de un modelo escalable de pagamento en cuanto el uso para sus iniciativas de inferencia de IA, forneciendo acceso a la expertise que acelera sus iniciativas.

El modelo OpEx reduce la inversión inicial y escala com sus necesidades de negocios em constante mudança, permitiendo que tú lleves perfectamente proyectos de la prova de concepto a la implementación y mucho más allá.



Implementación más rápida

Al substituir los requisitos de aprovação de *CapEx* y mudar para un modelo *OpEx*, el TruScale puede aumentar la flexibilidad y acelerar los tiempos de adquisición e implementación.



Opciones escalables

Escoge entre un contrato fijo o consumo medido para corresponder a las necesidad de su organización.



Expertise y servicios de IA integrados

Apoye-se en los servicios especializados de Lenovo para preencher lacunas de habilidades y recursos y ayudar a garantir el sucesso de la implementación. Além de isso, los gerentes de sucesso del cliente Lenovo dedicados pueden ayudar a facilitar y a coordinar com los recursos de Lenovo.

Esa flexibilidad no sólo torna más fácil para una gama más amplia de organizaciones alavancar la inferencia de IA, mas también torna la tecnología a la prova del futuro y elimina el riesgo de obsolescencia a medida que la tecnología evolui.





IA con un olho en la sustentabilidade

El aumento de la potencia computacional necesaria para entrenar y operar modelos de IA significa más eletricidade consumida y más calor gerado, lo que continua siendo una fuente de preocupación en todo el mundo.

Y, en cuanto la IA se integra a más aspectos de la vida cotidiana, el aumento resultante en la capacidade de la computación necesaria apenas acelerará.

Por ejemplo, una pesquisa típica del Google usa menos de 0,3 watt-hora (Wh) por solicitud. Adicionar una grande interacción de modelo de linguagem a la solicitud eleva ese requisito de energia para algo entre 7Wh y 9Wh para solicitud. Dado su volumen de pesquisa actual, si cada solicitud de pesquisa del Google incluyesse un componente de IA, la IA de Google sola podería consumir cerca de 30 terawatts-hora (TWh) para el año, o casi lo mismo, que el país de Irlanda⁶.



La demanda global de energia de IA está proyectada para aumentar **al menos 10x** até 2026.⁷

Lenovo e Intel están comprometidas com solucione sustentables, eficientes en termos de energia y ambientalmente responsable para la inferencia de IA.

Los processadores Intel® Xeon® 6 son los processadores de *data center* más sustentables de la Intel, entregando más de 2x major desempeño para *watt* en comparación con la generación anterior⁸. Y ellos pueden ser implementadas en *data centers* existentes sin requisitos adicionales de energia o resfriamiento.

En el *data center*, la tecnología de medición TruScale puede ayudá-lo a monitorar el consumo, la utilización y a temperatura para gerenciar el uso y los custos de manera más eficiente. Además de eso, nuestro software *Energy Aware Runtime* (EAR) y el *xClarity Energy Manager* ayudan a entregar el desempeño ideal en un bajo nivel de consumo de energia, otimizan los estados de energia, desligando componentes no utilizados y roteando cargas de trabajo para los recursos más eficientes.

Optimizar su data center con Lenovo TruScale ayuda a reducir las emisiones de CO₂ y el consumo de energia en hasta 20%⁹.



Pesquisa en el Google <0,3 Wh



Pesquisa no Google con tecnología de IA: 7-9 Wh



Todas las pesquisa no Google con IA: 30 TWh para el año

Una abordagem **más inteligente** para inferencia de IA en todos los lugares

La inferencia de IA detén una tremenda promesa de acelerar el crecimiento de los negocios, reducir las cargas de la trabajo y optimizar la eficiencia para empresas en todos los sectores.

No importa donde tú estés en la jornada para implementar soluciones basadas en IA en su organización, Lenovo e Intel están prontas para ayudar com soluciones construídas para fines específicos, *expertise* líder del y los mejores parceiros de la categoría.

Visita la página **Intel AI Alliance** para saber más.

Fuentes

- 1 CIO, "How AI is transforming business today," setiembre de 2024
- 2 Accenture, "Technology Vision 2023," marzo de 2023
- 3 Tidio, "10+ Essential AI Statistics You Need to Know for 2023," octubre de 2023
- 4 Datos de la Lenovo con base em pesquisa interna de la Lenovo ISG
- 5 Con base en testes internos de la Lenovo comparando com sistemas semejantes com resfriamento a ar en un data center típico
- 6 De Vries, "The growing energy footprint of artificial intelligence," Octubre de 2023
- 7 Agência Internacional de Energia, "Electricity 2024: Analysis and forecast to 2026," janeiro de 2024
- 8 Vea [9A2] en intel.com/processorclaims: Intel Xeon 6. Los resultados pueden variar.
- 9 TruScale IaaS gera relatórios precisos sobre consumo de energia y emisiones de CO2, permitindo que infraestructuras gerenciadas sean proyectadas, implementadas y ajustadas no apenas para desempeño y capacidad, mas también para emisiones de CO2. El monitoramiento continuo del sistema con Lenovo XClarity Power Monitor y dados de desempeño del sistemas es usado para optimización del consumo de energia de la infraestructura. Las emisiones de CO2 serán medidas a base de la pegada de carbono local de la fuente de energia utilizada.



Lenovo ThinkSystem SC750 V4 Neptune®, con procesador Intel® Xeon® 6 - Desempeño y eficiencia sin compromisos.

© Lenovo 2025. Todos los derechos reservados. v1.00 Enero 2025.

Intel, el logotipo de Intel, OpenVINO y el logotipo de OpenVINO son marcas comerciales de Intel Corporation o sus subsidiarias.

HOME

Smarter
technology
for all

Lenovo