



oneAPI

La Computación heterogénea

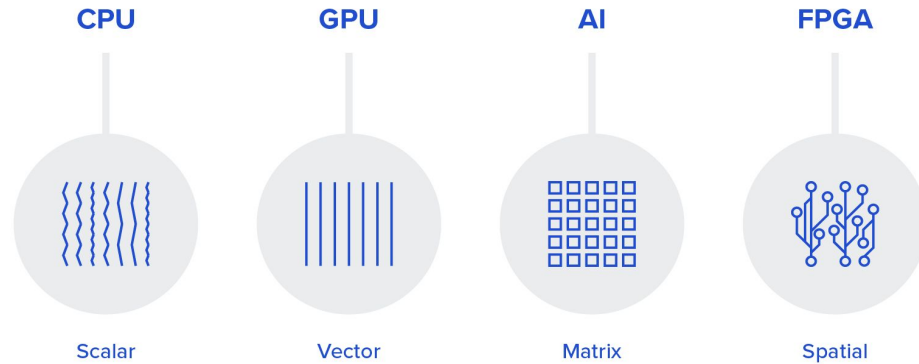
La **Computación heterogénea** se refiere a sistemas que usan más de un tipo de procesador.

Estos son sistemas que ganan en rendimiento no justo por añadir el mismo tipo de procesadores, sino por añadir procesadores distintos.

Normalmente, incorporan capacidades de procesamiento especializadas para realizar tareas particulares.

GPU (Unidad de Procesamiento Gráfico)	TPU (Unidad de Procesamiento Tensorial)	FPGA (Arreglo de Compuertas Programable en Campo)
Las GPU están diseñadas principalmente para tareas relacionadas con gráficos, como renderización de videojuegos, modelado 3D y cálculos científicos de alto rendimiento (por ejemplo, en aprendizaje profundo y aprendizaje automático).	Las TPU son aceleradores desarrollados por Google específicamente para el procesamiento de redes neuronales y tareas de aprendizaje profundo. Están optimizadas para operaciones matriciales y de tensor, lo que las hace ideales para tareas de IA.	Las FPGA son dispositivos lógicos reconfigurables que permiten a los desarrolladores crear hardware personalizado para tareas específicas. Pueden utilizarse en una amplia gama de aplicaciones, desde la criptografía hasta la aceleración de algoritmos de procesamiento de señales.

Arquitecturas involucradas



Escalar	Vectorial	Matricial	Espacial
cargas de trabajo complejas que funcionan mejor en una CPU .	cargas de trabajo que se pueden descomponer en vectores de instrucciones o vectores de elementos de datos y acelerar ejecutando el código en un procesador vectorial como una GPU .	cargas de trabajo, incluidas IA y ML , que realizan muchos cálculos matriciales y se ejecutan mejor en chips especializados de IA/ML.	cargas de trabajo que requieren procesadores únicos y especializados contruidos mejor sobre la marcha usando FPGA para satisfacer las necesidades computacionales de la carga de trabajo específica. Configurados adecuadamente, los FPGA pueden ejecutar cargas de trabajo escalares, vectoriales y matriciales, pero quizás no tan rápido o tan eficientemente como los procesadores especializados.

Intel oneAPI

Cada vez está más claro que el futuro de la informática no es un chip para todo, sino muchos chips para muchas cosas.

Pat Gelsinger, director ejecutivo de Intel, imaginó recientemente un “mar de aceleradores” entre los cuales los clientes podrían seleccionar, mezclar y combinar según sus necesidades específicas.

Suena bastante bien, ¿verdad? A menos que usted sea el desarrollador de software encargado de crear código personalizado para cada nuevo tipo de chip (o incluso simplemente probar un nuevo tipo de chip).

La gran idea detrás de oneAPI es permitir el uso de una plataforma para una variedad de hardware diferente, por lo que los desarrolladores no tendrían que usar diferentes lenguajes, herramientas y bibliotecas cuando codifican para CPU y GPU. Con oneAPI, la caja de herramientas y el código base serían los mismos para ambos.

Intel oneAPI

Data Parallel C++ (DPC++)

Data Parallel C++, es una alternativa de código abierto basada en C a los lenguajes de programación propietarios que normalmente se utilizan para codificar tipos específicos de hardware, como GPU o FPGAs. Data Parallel C++ también incorpora SYCL de Khronos Group para admitir el paralelismo de datos y la programación heterogénea.

Actualmente, Intel ofrece acceso gratuito a su DevCloud , lo que permite a los ingenieros de software probar sus herramientas y jugar con oneAPI y DPC++ en la nube sin muchos problemas.

¿OneAPI está restringida únicamente al hardware Intel?

No, oneAPI está diseñada para ser independiente del hardware y funcionar con CPU, GPU y varios aceleradores de hardware de diferentes proveedores.

¿OneAPI es de uso gratuito?

Sí, oneAPI es una iniciativa de código abierto y casi todos sus componentes básicos se pueden utilizar libremente.

Intel oneAPI

Application Workloads Need Diverse Hardware

Middleware & Frameworks



1
oneAPI Intel®oneAPI
Product

Compatibility
Tool

Languages

Libraries

Analysis &
Debug Tools

Low-Level Hardware Interface

XPU's



CPU



GPU



FPGA

Intel oneAPI Toolkits

Cajas de herramientas (Toolkits) especializadas para satisfacer las necesidades de diferentes áreas de desarrollo.

1. Intel oneAPI Base Toolkit: Este es el conjunto de herramientas base que proporciona compiladores, bibliotecas y utilidades esenciales para el desarrollo de software en una variedad de dominios.

2. Intel oneAPI HPC Toolkit : Diseñado para la informática de alto rendimiento (HPC), incluye herramientas y bibliotecas para la programación paralela y distribuida, así como optimizaciones para aceleradores y procesadores Intel.

3. Intel oneAPI AI Analytics Toolkit : Orientado a la inteligencia artificial (IA) y análisis de datos, proporciona herramientas y bibliotecas para el desarrollo de aplicaciones de aprendizaje profundo y análisis de datos de alto rendimiento.

4. Intel oneAPI IoT Toolkit : Centrado en el Internet de las cosas (IoT), este toolkit ofrece herramientas para el desarrollo de aplicaciones IoT que pueden aprovechar el rendimiento y la eficiencia energética de los procesadores Intel.

5. Intel oneAPI Rendering Toolkit : Diseñado para gráficos y visualización, este toolkit proporciona bibliotecas y herramientas para el desarrollo de aplicaciones de renderización y simulación.

6. Intel oneAPI Threading Building Blocks (TBB) : Aunque no es un toolkit independiente, TBB es una biblioteca importante para la programación paralela y de múltiples hilos que se utiliza en combinación con otros toolkits.

Intel oneAPI AI Analytics toolkit

What's Inside: Intel® oneAPI AI Analytics Toolkit

DEEP LEARNING

Intel® Extension for TensorFlow

Intel® Extension for PyTorch

Model Zoo for Intel® Architecture

Intel® Neural Compressor

MACHINE LEARNING

Intel® Extension for Scikit-learn

Intel-Optimized XGBoost

DATA ANALYTICS

Intel® Distribution of Modin

OmniSci Backend

INTEL-OPTIMIZED PYTHON

NumPy

SciPy

Numba

Pandas

Data Parallel Python



CPU



FUTURE GPU ACCELERATORS

Supported Hardware Architecture (Hardware support varies by individual tool.
Architecture support will be expanded over time.)

Fuente de las mejoras

Intel Extensión para Scikit-learn utiliza la biblioteca de análisis de datos Intel oneAPI (oneDAL) para lograr su aceleración. Esta biblioteca habilita todas las instrucciones vectoriales más recientes, como Intel Advanced Vector Extensions (Intel AVX-512). También utiliza bloqueo de datos compatible con caché, operaciones BLAS rápidas con Intel oneAPI Math Kernel Library (oneMKL) y subprocesos múltiples escalables con Intel oneAPI Threading Building Blocks (oneTBB).

- oneAPI Data Analytics Library (oneDAL) es una potente biblioteca de aprendizaje automático que le ayuda a acelerar el análisis de big data en todas las etapas: preprocesamiento, transformación, análisis, modelado, validación y toma de decisiones.
- La biblioteca implementa algoritmos clásicos de aprendizaje automático. El aumento de su rendimiento se logra aprovechando las capacidades del hardware Intel®.
- oneDAL es parte de oneAPI. La rama actual implementa la versión 1.1 de la especificación oneAPI.

Intel® Extensión para Scikit-learn*

Acelera sin problemas sus aplicaciones scikit-learn para CPU y GPU Intel en configuraciones de uno y varios nodos.

- | | |
|--|---|
| <ul style="list-style-type: none">• Acelerar el entrenamiento y la inferencia hasta 100 veces con la precisión matemática equivalente. | <ul style="list-style-type: none">• Continuar utilizando la API scikit-learn de código abierto.• Intel Extensión para Scikit-learn admite sistemas Linux, Windows y Mac en arquitecturas x86. Se puede descargar usando PyPI o Anaconda Cloud. |
| <ul style="list-style-type: none">• Habilitar y deshabilitar la extensión con un par de líneas de código o en la línea de comando. | <ul style="list-style-type: none">• Acelere los algoritmos de scikit-learn (sklearn) reemplazando los estimadores existentes con versiones aceleradas matemáticamente equivalentes (ver algoritmos soportados). |

Este paquete de extensión parchea dinámicamente los estimadores de scikit-learn mientras mejora el rendimiento de sus algoritmos de aprendizaje automático.

Elige cómo aplicar las aceleraciones:

Parchee todos los algoritmos compatibles desde la línea de comando sin cambios de código.

Agregue dos líneas de código para parchear todos los algoritmos compatibles en su secuencia de comandos Python.

Parchee y elimine parches globalmente de su entorno para todos los usos de scikit-learn.

Especifique en su script parchear solo los algoritmos seleccionados.

Validando Intel® Extensión para Scikit-learn* en Google Colaboratory

Publicado en

[https://github.com/anissval/oneAPI/blob/main/Trabajo%20Final%20-%20OneAPI%20-%20Anahi%20Valdiviezo/Demostracion%20uso%20de%20Intel%20AI%20Analytics%20Toolkit/Google Colaboratory random forest yolanda.ipynb.pdf](https://github.com/anissval/oneAPI/blob/main/Trabajo%20Final%20-%20OneAPI%20-%20Anahi%20Valdiviezo/Demostracion%20uso%20de%20Intel%20AI%20Analytics%20Toolkit/Google%20Colaboratory%20random%20forest%20yolanda.ipynb.pdf)

Resultados de la comparación de la métrica MSE de Scikit-learn parcheado y original

Métrica MSE de Scikit-learn parcheado: 83.60018104634611 -> **mejor modelo**

Métrica MSE de Scikit-learn sin parches: 83.80131297814816

Relación de métricas: 0,9975998952205618

Original Scikit-learn time: **4453.85 s**

Intel® extension for Scikit-learn time: 2081.74 s

Obtén aceleración en 2,1 veces.



La CPU predeterminada para Google Colaboratory es una CPU Intel Xeon con 2 vCPU (CPU virtuales) y 13 GB de RAM.

Ejemplo Algoritmos soportados

Consulte la lista de algoritmos y parámetros admitidos para obtener más detalles

(<https://intel.github.io/scikit-learn-intelx/latest/algorithms.html>).

En los casos en que se utilizan parámetros no compatibles, el paquete recurre al Scikit-learn original. Si el parche no cubre sus escenarios, envíe un problema en GitHub.

Supported Algorithms

Applying Intel® Extension for Scikit-learn* impacts the following scikit-learn algorithms:

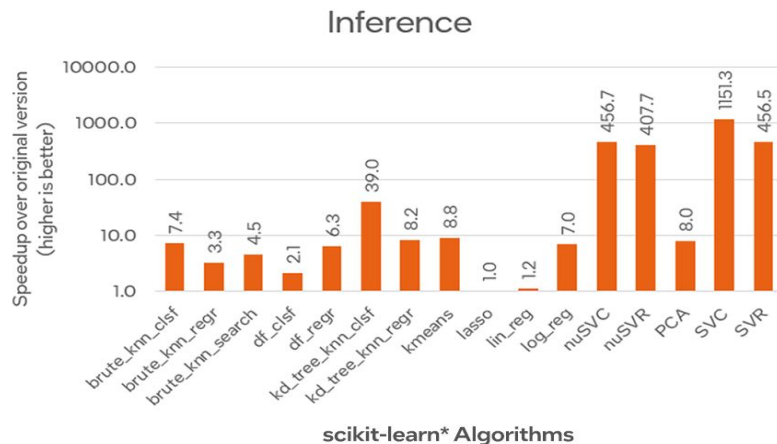
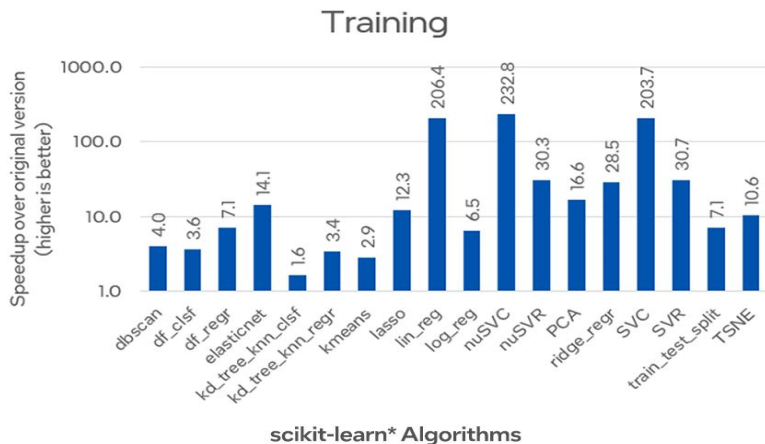
on CPU

Classification

Algorithm	Parameters	Data formats
SVC	All parameters are supported	No limitations
NuSVC	All parameters are supported	No limitations
<i>RandomForestClassifier</i>	All parameters are supported except: <ul style="list-style-type: none"><code>warm_start</code> <code>!= True</code><code>cpp_alpha</code> <code>!= 0</code><code>criterion</code> <code>!= 'gini'</code>	Multi-output and sparse data are not supported
<i>KNeighborsClassifier</i>	<ul style="list-style-type: none">For <code>algorithm == 'kd_tree'</code>: all parameters except <code>metric</code> <code>!= 'euclidean'</code> or <code>'minkowski'</code> with <code>p</code> <code>!= 2</code>For <code>algorithm == 'brute'</code>: all parameters except <code>metric</code> not in <code>['euclidean', 'manhattan', 'minkowski', 'chebyshev', 'cosine']</code>	Multi-output and sparse data are not supported
<i>LogisticRegression</i>	All parameters are supported except: <ul style="list-style-type: none"><code>solver</code> not in <code>['lbfgs', 'newton-cg']</code><code>class_weight</code> <code>!= None</code><code>sample_weight</code> <code>!= None</code>	Only dense data is supported

Performance Benefits of Intel Extension for Scikit-learn*

Performance Benefits of Intel® Extension for Scikit-learn* Combined Averages of FP32 & FP64 Workloads



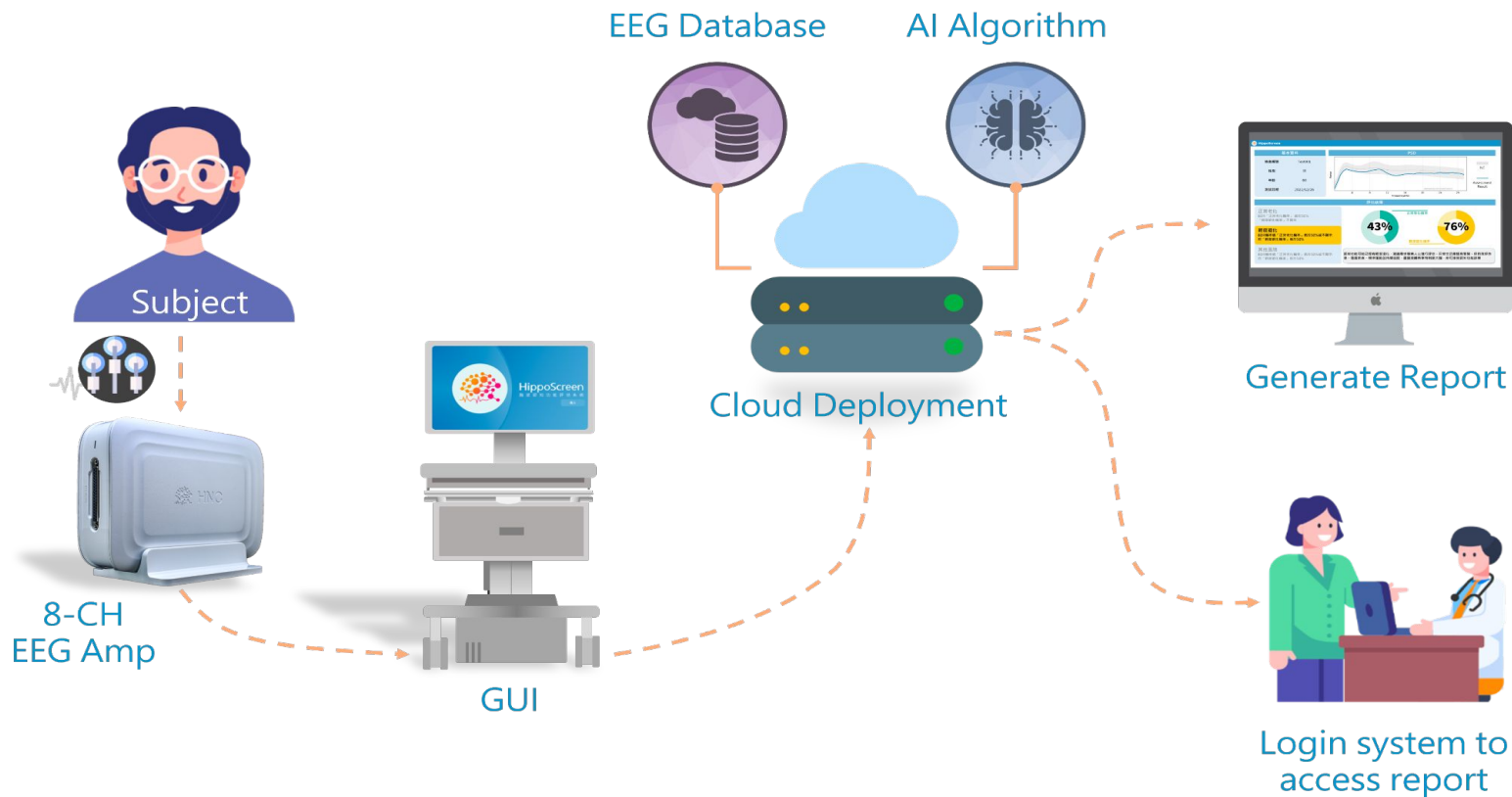
Testing Date: Performance results are based on testing by Intel as of March 21, 2023 and may not reflect all publicly available security updates.

Configuration Details and Workload Setup: bare metal (2.0 GHz Intel® Xeon® Platinum 8480+ processor, two sockets, 56 cores per socket), 512 GB DDR5 4800MT/s, Python® 3.10, scikit-learn 1.2.0, scikit-learn-intelex 2023.0.1. Intel optimizations include use of multithreading implementation for scikit-learn algorithms (which are typically single-threaded), as well as other hardware and software optimizations.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more at www.intel.com/PerformanceIndex. Your costs and results may vary.

HippoScreen y oneAPI mejoran las pruebas de salud mental



HippoScreen y oneAPI mejoran las pruebas de salud mental

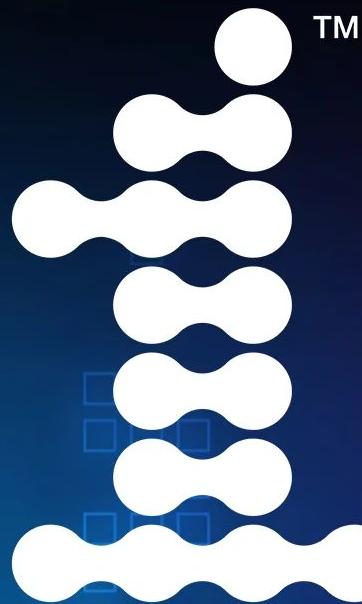
HippoScreen aprovecha las herramientas Intel oneAPI y los frameworks de IA para optimizar los modelos de deep-learning que ayudan a diagnosticar la depresión.

Utilizando el Intel® AI Analytics Toolkit y el Intel® oneAPI Base Toolkit, HippoScreen mejoró la eficiencia y los tiempos de construcción de los modelos de deep-learning utilizados en su sistema de inteligencia artificial (IA) de ondas cerebrales. Estas mejoras permitieron a HippoScreen ampliar las aplicaciones de su sistema a una gama más amplia de afecciones y enfermedades psiquiátricas.



"En HippoScreen hemos podido aprovechar las optimizaciones de software de Intel® Extension for Scikit-learn e Intel® Optimization for PyTorch para acelerar 2,4 veces los tiempos de compilación de los modelos de IA en nuestro sistema personalizado de análisis de ondas cerebrales EEG", afirma Daniel Weng, director de tecnología de HippoScreen NeuroTech Corp.

Gracias!



oneAPI