FUNDAMENTALS OF ACCELERATED COMPUTING WITH CUDA PYTHON

NVIDIA | DEEP LEARNING INSTITUTE

This workshop teaches you the fundamental tools and techniques for running GPU-accelerated Python applications using CUDA® GPUs and the Numba compiler. You'll work though dozens of hands-on coding exercises and, at the end of the training, implement a new workflow to accelerate a fully functional linear algebra program originally designed for CPUs, observing impressive performance gains. After the workshop ends, you'll have additional resources to help you create new GPU-accelerated applications on your own.

Learning Objectives

At the conclusion of the workshop, you'll have an understanding of the fundamental tools and techniques for GPU-accelerated Python applications with CUDA and Numba:

> GPU-accelerate NumPy ufuncs with a few lines of code.
> Configure code parallelization using the CUDA thread hierarchy.
> Write custom CUDA device kernels for maximum performance and flexibility.
> Use memory coalescing and on-device shared memory to increase CUDA kernel bandwidth.

Workshop Information and Prerequisites:

Duration:	8 hours
Price:	**Contact us for pricing.**
Prerequisites:	Basic Python competency, including familiarity with variable types, loops, conditional statements, functions, and array manipulations. NumPy competency, including the use of ndarrays and ufuncs. No previous knowledge of CUDA programming is required.
Tools, libraries, and frameworks:	Numba, NumPy
Assessment type:	Code-based
Certificate:	Upon successful completion of the assessment, participants will receive an NVIDIA DLI certificate to recognize their subject matter competency and support professional career growth.
Hardware/software requirements:	Desktop or laptop computer capable of running the latest version of Chrome or Firefox. Each participant will be provided with dedicated access to a fully configured, GPU-accelerated workstation in the cloud.
Languages:	English

Sample Workshop Outline

Introduction (15 mins)	> Meet the instructor. > Create an account at **courses.nvidia.com/join**
Introduction to CUDA Python with Numba (120 mins)	> Begin working with the Numba compiler and CUDA programming in Python. > Use Numba decorators to GPU-accelerate numerical Python functions. > Optimize host-to-device and device-to-host memory transfers.
Break (60 mins)	
Custom CUDA Kernels in Python with Numba (120 mins)	> Learn CUDA's parallel thread hierarchy and how to extend parallel program possibilities. > Launch massively parallel, custom CUDA kernels on the GPU. > Utilize CUDA atomic operations to avoid race conditions during parallel execution.
Break (15 mins)	
Multidimensional Grids, and Shared Memory for CUDA Python with Numba (120 mins)	> Learn multidimensional grid creation and how to work in parallel on 2D matrices. > Leverage on-device shared memory to promote memory coalescing while reshaping 2D matrices.
Final Review (15 mins)	> Review key learnings and wrap up questions. > Complete the assessment to earn a certificate. > Take the workshop survey

Why Choose NVIDIA Deep Learning Institute for Hands-On Training?

> Access workshops from anywhere with just your desktop/laptop and an internet connection. Each participant will have access to a fully configured, GPU-accelerated workstation in the cloud.

> Obtain hands-on experience with the most widely used, industry-standard software, tools, and frameworks.

> Learn to build deep learning and accelerated computing applications for industries, such as healthcare, robotics, manufacturing, accelerated computing, and more.

> Gain real-world experience through content designed in collaboration with industry leaders, such as the Children's Hospital of Los Angeles, Mayo Clinic, and PwC.

> Earn an NVIDIA Deep Learning Institute certificate to demonstrate your subject matter competency and support your career growth.

For the latest DLI workshops and trainings, visit **www.nvidia.com/dli**

For questions, contact us at **nvdli@nvidia.com**