



Fundamentals of Deep Learning for Multi-GPUs

This workshop teaches you to apply techniques to train deep neural networks on multiple GPUs to shorten the training time required for data-intensive applications. You will work with widely-used deep learning tools, frameworks, and workflows by performing neural network training on a fully-configured GPU accelerated workstation in the cloud. The workshop starts with a linear neuron, defining the loss function and optimization logic for gradient descent. It then teaches the concepts to transform single GPU to Horovod multi-GPU implementation to reduce the complexity of writing efficient distributed software and then concludes by teaching the techniques to improve the overall performance of the whole pipeline.

Duration	8 hours
Price	\$10000 for groups of up to 20 people (includes dedicated access during the course to a fully-configured GPU accelerated workstation in the cloud for each student)
Certification	Yes
Prerequisites	Experience with Stochastic Gradient Descent
Languages	English
Tools, Libraries, and Frameworks	TensorFlow

Learning Objectives

At the conclusion of the workshop, you will have an understanding of:

- Various approaches to multi GPU training
- Algorithmic and engineering challenges to the large-scale training of a neural network

Why Deep Learning Institute Hands-on Training?

- Learn how to build deep learning and accelerated computing applications across a wide range of industry segments such as Autonomous Vehicles, Digital Content Creation, Finance, Game Development, and Healthcare
- Obtain guided hands-on experience using the most widely used, industry-standard software, tools, and frameworks
- Attain real world expertise through content designed in collaboration with industry leaders such as the Children’s Hospital of Los Angeles, Mayo Clinic, and PwC
- Earn NVIDIA DLI Certification to prove your subject matter competency and support professional career growth
- Access content anywhere, anytime with a fully configured GPU-accelerated workstation in the cloud



Fundamentals of Deep Learning for Multi-GPUs

Content Outline

	Components	Description
Theory of Data Parallelism (45 mins)	<ul style="list-style-type: none"> Issues with sequential processing 	Understand the issues with sequential single thread data processing and speeding up the applications with parallel processing.
Break (15 mins)		
Introduction to Multi GPU training (120 mins)	<ul style="list-style-type: none"> Overview of loss function, gradient descent, and stochastic gradient descent(SGD)s 	Define a simple neural network and a cost function and iteratively calculate the gradient of the cost function and model parameters using the SGD optimization algorithm.
Break (60 mins)		
Algorithmic challenges to Multi GPU training (120 mins)	<ul style="list-style-type: none"> Data parallelism Large minibatch and its impact on accuracy Gradient exchange 	Learn to transform single GPU to Horovod multi-GPU implementation to reduce the complexity of writing efficient distributed software. Understand the data loading, augmentation, and training logic using AlexNet model.
Break (15 mins)		
Engineering challenges to Multi GPU training (120 mins)	<ul style="list-style-type: none"> Keeping up with the GPU Job Scheduling Overview of the wider AI system design 	Understand the aspects of data input pipeline, communication, reference architecture and take a deeper dive into the concepts of job scheduling.
Closing Comments & Questions (15 mins)	<ul style="list-style-type: none"> Wrap-up with the potential next steps and Q&A 	Quick overview of the next -steps you could leverage to build and deploy your own applications and any Q&A