Contribution ID: **2**   Type: **not specified**

Stochastic Gradient Descent and the Effects of Batch Size

Monday, November 6, 2023 9:15 AM (2 hours)

Learn the significance of stochastic gradient descent when training on multiple GPUs

- Understand the issues with sequential single-thread data processing and the theory behind speeding up applications with parallel processing.
- Understand loss function, gradient descent, and stochastic gradient descent (SGD).
- Understand the effect of batch size on accuracy and training time with an eye towards its use on multi-GPU systems.