

Part 2/3:

Memory and ML Model Execution

- 9.** Memory Hierarchy - RAM, VRAM, Cache
- 10.** ML Models Loading and Memory Management
- 11.** Batch Size and Memory Usage Calculation
- 12.** Python Bytecode and ML Frameworks
- 13.** Memory Layout of NumPy/PyTorch Tensors
- 14.** GPU Memory Management
- 15.** Mixed Precision Training - FP16 and FP32
- 16.** Hands-on: Resource Monitoring Tools