UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

Seminar Announcement

DIPARTIMENTO DI INFORMATICA

The Seminars on "Information Technology Outlook" – PhD Program in Computer Science and Mathematics

Dott. Giuseppe FIAMENI

Data Scientist - NVIDIA AI Technology Center - Italy

Lunedì 19 giugno 2023, ore 10:30 – Sala Consiglio, VII piano Dipartimento di Informatica

Large-scale model training with GPUs

Over the past decade, Deep Learning has emerged as the most significant breakthrough in computer science, particularly when it comes to large-scale training. It has revolutionized prediction models across various research topics and application fields, including computer vision, natural language processing, embodied AI, and traditional pattern recognition. The effectiveness of deep learning models has significantly advanced, thanks to their ability to handle complex problems with vast amounts of data. As deep learning models grow in complexity to tackle increasingly challenging problems, the demand for scalable methods and software to train them has also grown. The computational approach has shifted from CPU-only methods to leveraging the power of GPUs and other massively parallel devices. These hardware advancements, coupled with the availability of large-scale and highly dimensional datasets, have facilitated the training of deep learning models on a much larger scale. The objective of this tutorial is to provide attendees with a comprehensive understanding of deep learning on high-performance computing (HPC)-class systems, with a specific focus on large-scale training. Participants will gain practical knowledge of core concepts, performance optimization techniques, and valuable tips and techniques for scaling deep learning models to handle massive datasets.

Giuseppe Fiameni is a Data Scientist at NVIDIA where he oversees the NVIDIA AI Technology Center in Italy, a collaboration among NVIDIA, CINI and CINECA to accelerate academic research in the field of Artificial Intelligence through collaboration projects. He has been working as HPC specialist at CINECA, the largest HPC facility in Italy, for more than 14 years providing support for large-scale data analytics workloads.