CS5463: High Performance Machine Learning

Spring 2026

Term Project Write Up

Topic: “Comparative Analysis of Distributed Deep Learning Frameworks”

Submitted By:
Anshu Tripathi (wbb206)

Problem Statement:

The rapid advancement of deep learning has made multi-GPU and distributed systems essential for efficient model training and large-scale experimentation. Modern distributed deep learning frameworks employ different synchronization strategies, communication backends, and optimization techniques that significantly impact scalability and overall performance.

Although several frameworks are widely used, their relative efficiency, communication overhead, and training scalability on high-performance computing (HPC) systems require systematic evaluation. Understanding these trade-offs is critical for selecting appropriate distributed training strategies in practical HPC environments.

Objective:

This project aims to perform a systematic comparative analysis of three distributed deep learning training approaches on a multi-GPU platform:

1. PyTorch Distributed Data Parallel (DDP)
2. Horovod-based distributed training
3. PyTorch Distributed Data Parallel with Automatic Mixed Precision (AMP)

The experiments will be conducted on the UTSA ARC cluster using GPUs. A standard image classification task will be used to ensure controlled and reproducible evaluation. The comparison will focus on:

- Training time per epoch
- Scalability as the number of GPUs increases
- Speedup and efficiency
- Communication overhead
- GPU memory utilization
- Final model accuracy and convergence behavior

The underlying communication mechanisms will be analyzed to understand synchronization patterns and performance bottlenecks. Additionally, the impact of mixed precision training on computational throughput and memory efficiency will be evaluated.

This study aims to provide practical insights into the performance trade-offs among distributed deep learning frameworks and identify which approach achieves the best balance between scalability, efficiency, and accuracy on HPC systems.

Project Repository:

All source code, documentation, experiment logs, and performance plots will be maintained in the following GitHub repository:

https://github.com/anshutripathi11/HPML_Project.git