FLOPS and ML Model Performance Metrics

FLOPS

Floating Point Operations Per Second

Hardware capability metric

FLOPs

Floating Point Operations

Model complexity metric (total operations required)

GPT-3 Training Example

Total FLOPs: **~3.14×10^{23}** (314 zettaFLOPs)

GPU Performance Comparison

NVIDIA A100

312

TFLOPS

NVIDIA H100

3.2x faster

1000

TFLOPS

Training Time Calculation

Training Time = **Model FLOPs** / **Hardware FLOPS** / **GPU Count**