

Memory Reduction through 4-bit Quantization

