

Best Learning Rate vs Compute (Non-embedding)
ademamix | eryngii_sweeps

