

Adam Denominator vs Batch-Size LR Scaling

