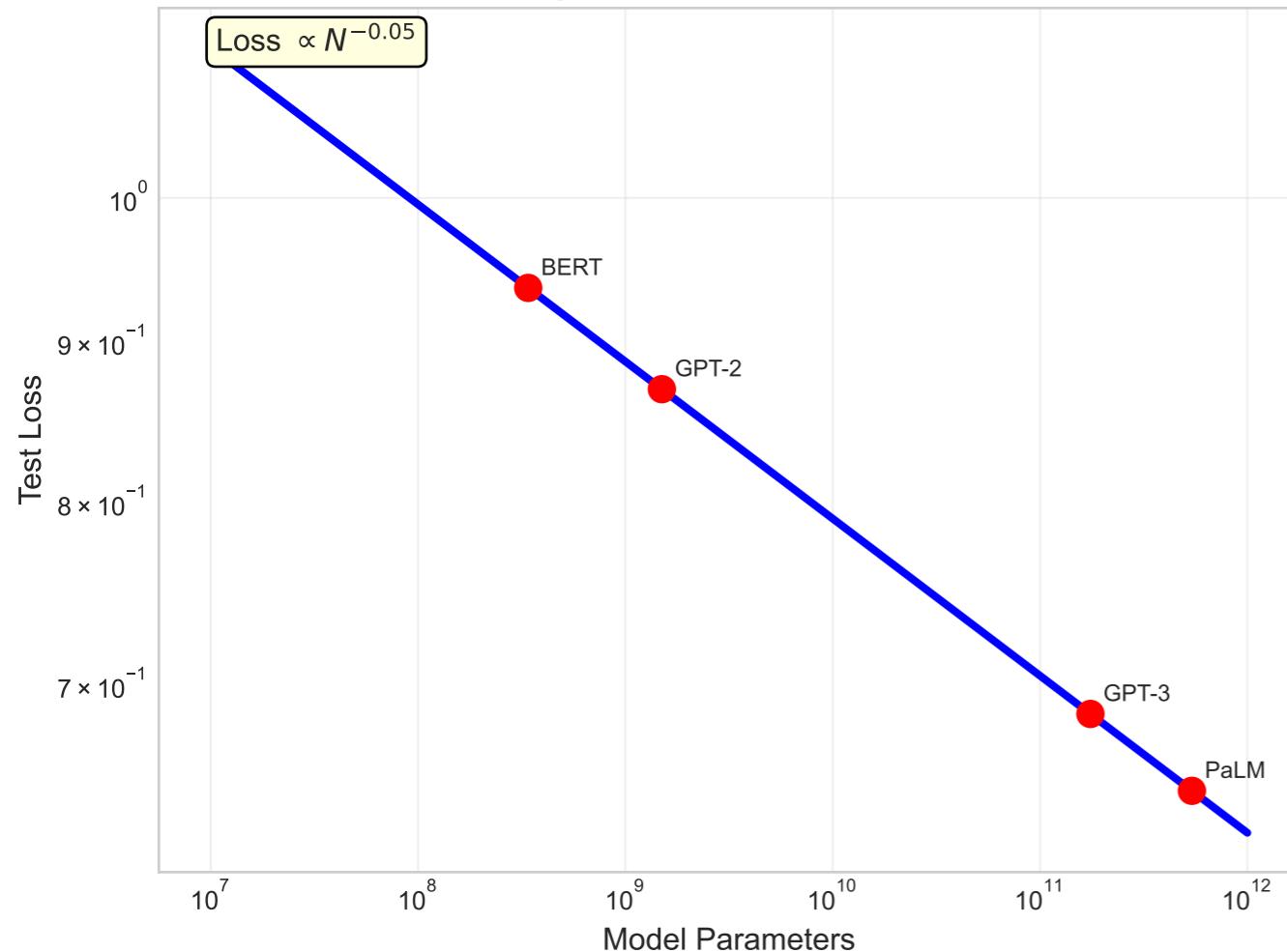


Scaling Laws: Predictable Progress with Scale

Scaling Law: Loss vs Parameters



Chinchilla Scaling: Compute-Optimal Training

