

スケーリング則の活用

Chinchillaの法則による最適化:

- 従来の考え方: より大きいモデル = より良い性能
- 新しい発見: データとモデルサイズのバランスが重要

具体例:

- GPT-3 (175B/1Tトークン) vs Chinchilla (70B/4Tトークン)
 - 同じ計算コストでより高い性能
 - 4倍のデータ、1/2.5のパラメータ数

実践的な利点:

- 計算効率の向上
- コスト削減
- 環境負荷の低減