

事前学習済み言語モデル (Pre-trained Language Models)

- BERTは、TransformerアーキテクチャとSelf-attentionメカニズムに基づき、大規模なコーパス上を使って双方向言語モデルを事前学習することによって提案された
- 事前訓練された文脈を考慮した単語表現は、汎用的な意味要素として非常に効果的であった
- 事前学習と微調整という考え方は、のちの研究に影響を与えた

大規模言語モデル (Large Language Models)

- より大きなPLM
 - ➡ PLMのスケーリングは、多くの場合下流タスクでのモデル性能の向上につながる事がわかっている（スケール則に従う）
- LLMとPLMには3つの大きな違いがある
 1. 創発的能力がみられ、ある程度汎用の言語タスクソルバーとして機能する
 - ➡ 汎用人工知能 (AGI) の可能性が再考されている
 2. 人間がAIを開発し、利用する方法に革命をもたらす
 - ➡ LLMがどのように動作するか理解し、従わせるようにタスクを作る必要性
 3. LLMの開発は、研究とエンジニアリングの明確な区別がない
 - ➡ 大規模なデータ処理や並列分散処理などの実践的な経験が必要