

GPT-3

- 2020年にリリースされた175Bのパラメータを持つモデル
- 文脈内学習 (In-context learning) の概念が正式に導入された
- OpenAIはこのモデルをさらに改良するために2つの手法をとった

1. コードデータでのトレーニング

コードベースのモデルを元にトレーニングしたGPT-3.5は推論能力が大幅に向上した

2. 人間によるアライメント

人間のフィードバックを組み込むことで命令追従能力が向上し、有害なコンテンツの生成を軽減できた

ChatGPT

- 2022年11月にリリースされたGPTモデルをベースにした会話モデル
- 人間が生成した会話 (ユーザ役とAI役の両方を演じたもの) を使用してトレーニング
- 膨大な知識の蓄積・推論スキル・文脈の保持・人間の価値観への適合により、人間とのコミュニケーションで優れた能力を示した
- プラグインなどを組み合わせればさらに能力を拡張できる