

Pythonで作って学ぶ LLM (大規模言語モデル)

石垣 達也

● 線形代数との関係

LLM (Large Language Model: 大規模言語モデル) はChatGPTなどのテキスト生成AIの基盤技術です。テキスト生成AIは、ユーザの入力する指示 (プロンプト) に対し、例えば、「英語の〇〇を日本語に翻訳して」、「次の新聞記事を要約して」というプロンプトを与えると、翻訳や要約といった知的な処理を行います。

テキスト生成AIすなわち、LLMの背後では、トランス

フォーマ言語モデルと呼ばれる数理モデルがプロンプトを読み込み、返答を生成しています。この数理モデルは具体的には、人間の言葉を数学の世界のベクトルとして表現し、人間が思考する過程を行列演算の繰り返しとして表現する興味深い処理を行います。

本稿は、LLMの基盤技術であるトランスフォーマ言語モデルをPythonで実際に実装することで、LLMの知的な処理について、より深い理解を目指します。

3-1 仕組み…トランスフォーマという数理モデルを使う

● LLMとは言語用に特化した生成AI

LLMとは何かを、LLMとして有名なモデルである、

ChatGPTを例に簡単に説明します。ChatGPTはGenerative Pre-trained Transformer (GPT) と呼ばれ

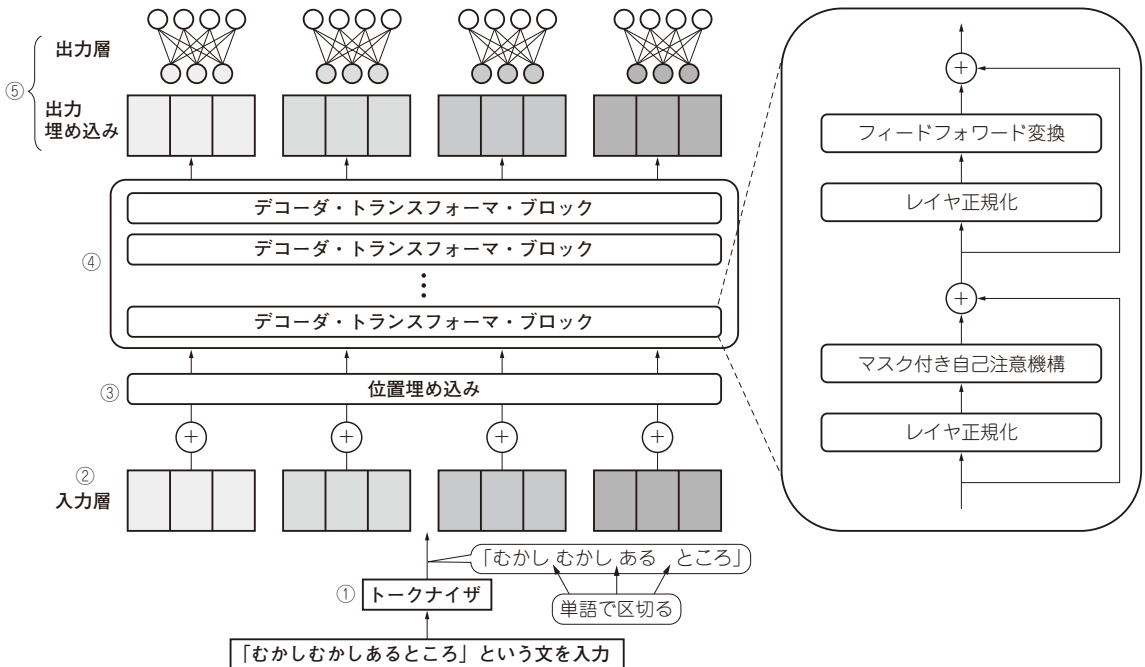


図1 (1) トランスフォーマの概要

文献(1)を元に筆者が加筆