

連載

放射線科では AI Chatは こう使う!

第1回の記事では、AI Chatの多様性とその広範な用途について、伊藤倫太郎先生から豊富な情報をご提供いただきました。そこで今回は、特定の使い方を1つ取り上げ、それを少し掘り下げてみることにしましょう。特に、データ解析に活用する方法に焦点を当ててみます。

第2回

中村 優太

東京大学医学部附属病院 22世紀医療センター
コンピュータ画像診断学／予防医学講座

読者の皆さんには、日頃から研究を精力的にされている方もいらっしゃるやれば、研究にはそれほど馴染みはないけれども、業務改善の一環として業務データの解析を任せられたという方もいらっしゃるかもしれません。

一般的に、データ解析には以下のようなソフトウェアやプログラミング言語がよく用いられることと思います：

- ・主にクリックで進めて行くもの：Excel、EZR、SPSS、JMPなど
- ・主にコーディングで進めて行くもの：R、Python、SAS、STATAなど

これらは強力なツールですが、習熟するまでには時間がかかります。また熟練者であっても、複雑なデータ操作を組み合わせたり長いソースコードを書いたりする必要があり、時間がかかることがあります。

ここで生成AIの出番です。

生成AIにとって、ソースコードの生成は得意な部類のタスクに入ります。このため生成AIを用いると、コーディングの手間を省き、データ解析を効率よく進めることができます。

たとえば、図1のようなダミーデータを考えてみましょう。仮に私達がいま、ある疾患の画像から自動で重症度を判定するシステムを開発していて、その性能評価実験の結果が得られたところであるとします。このデータは表形式で、以下の情報を含んでいるものとします：



- ・ID：仮名化された患者ID。
- ・ground_truth：重症度を示す正解データ。順序尺度で、0～3のいずれかの整数をとる。
- ・human_1、human_2：2名の医療従事者による重症度の判定。
- ・system_1、system_2：2つの自動判定システムによる重症度の判定。

このダミーデータを題材にして、生成AIにデータ解析を手伝ってもらいましょう。

データ解析にChatGPTを使う

1. 安全に解析するには？

ChatGPTにデータ解析を手伝ってもらう場合、データ自体を直接ChatGPTに渡す方法と、渡さない方法の2通りが存在します。

データ自体を直接ChatGPTに渡す方法の一つは、ChatGPTの有料版（GPT-4）のデータ解析機能“Advanced Data Analysis”を用いる方法です。これは、言葉で指示するだけで、データ解析やグラフ作成などが自動的に行われ、すべてをChatGPT上で完結できる大変魅力的な機能です。

しかし、医療機関で取得したデータには、要配慮個人情報が