

## 医学科第3学年「研究室配属」報告書

提出日 2025 年 10 月 24 日

【学籍番号】	112347
【学生氏名】	瀧本健一
【配属先講座等】	マルチメディアセンター
【指導教員名】	本山一隆
【レポートテーマ】	大規模言語モデルGeminiにおけるtemperature 設定が医師国家試験の解答の精度に与える影響

## 概要

本研究では、大規模言語モデル Gemini を用い、temperature 設定の違いが医師国家試験問題に対する解答性能に与える影響を検討した。第 118 回の医師国家試験問題の A 問題の PDF から問題文を抽出し、temperature を 0.2, 0.4, 0.6, 0.8, 1.0 の各値に設定して Gemini に回答させた。また、プロンプトは 2 種類用意した。各回答について正答率および出力内容の一貫性を評価した結果、temperature が低い条件では回答の安定性が高かった。一方、temperature を高く設定すると、回答のフォーマットに多様性がみられるようになったものの、正答率に大きな変化はみられなかった。これらの結果は、医師国家試験レベルの問題において、Gemini の回答が temperature 設定に対して比較的頑健であることを示唆するものである。

## 1. 序論

近年、大規模言語モデル（Large Language Model, LLM）の発展により、人工知能が自然言語処理だけでなく、教育・医療分野での応用も注目されている。特に、医師国家試験のような高難度の専門知識を要する問題に対して、LLM がどの程度正確に解答できるかは、AI の教育的利用や学習支援の可能性を評価する上で重要な指標となる。

LLM の出力は、生成アルゴリズムのパラメータによって変動することが知られている。その中でも、temperature は生成のランダム性や多様性を調整する主要なハイパーパラメータであり、低い値では決定的な出力、高い値では多様な応答が生成されるとされる。

本研究では、LLM の一つである Gemini を対象に、複数の temperature 設定における医師国家試験問題への回答精度を比較し、出力の安定性や正答率への影響を評価することを目的とする。

## 2. 方法

本研究では、大規模言語モデル Gemini-2.0-flash を用いて、医師国家試験問題への回答性能を評価した。実験環境は Google Colab 上で構築し、Python を用いて API 経由でモデルを呼び出した。

対象とした問題は、厚生労働省が公表している医師国家試験の A 問題 PDF から、Colab 上で文字列として抽出した。抽出には Python の PDF 解析ライブラリを使用し、選択肢と問題文を適切に整形した。

プロンプトは 2 種類用意し、各プロンプトに対して temperature を 0.2 から 1.0 まで、0.2 刻みで設定した。各条件において、モデルは 5 回ずつ回答を生成した。これにより、temperature とプロンプトの組み合わせごとの回答傾向や正答率の変動を比較可能とした。

実験に用いたソースコードは GitHub 上で公開しており、以下のリンクから参照可能である：<https://github.com/ユーザー名/リポジトリ名>

必要に応じて、本レポート中にもソースコードの一部を掲載し、モデル呼び出し例やプロンプトの設定方法を示した。例えば、Python によるモデル呼び出しは以下のように記述した：

```
import openai

response = openai.chat.completions.create(
    model="gemini-2.0-flash",
    messages=[{"role": "user", "content": prompt}],
```

temperature=temp

)

このようにして，temperature の異なる条件で得られた回答を収集し，正答率および出力内容の傾向を比較分析した．

### 3. 結果

図や表を用いて結果を明示する（例：図??）．

### 4. 考察

結果の解釈を述べる．先行研究との比較，考えられる機序，研究の限界などを整理する．

### 5. 結論

研究の主な結論を 1～3 文程度で明確に述べる．

### 謝辞

指導教員や協力者への謝辞を簡潔に記す．

### 参考文献

[1] 山田太郎ほか. ○○に関する研究. 日本医学会誌.2023; 112(4): 123–130.

[2] Smith J, et al. Analysis of ○○. J Med Sci. 2022; 45(2): 200–210.