

A JMMLU Tasks

JMMLU consists of 7,536 questions in the following 56 tasks (subjects). All tasks and their numbers are shown in Table 3.

| Task Name | Number | Task Name | Number |
|---------------------------------|--------|---|--------|
| 専門医学 (professional_medicine) | 150 | 高校心理学 (high_school_psychology) | 150 |
| 専門心理学 (professional_psychology) | 150 | 高校物理 (high_school_physics) | 150 |
| 専門会計 (professional_accounting) | 150 | 高校統計学 (high_school_statistics) | 150 |
| 哲学 (philosophy) | 150 | 高校数学 (high_school_mathematics) | 150 |
| 雑学 (miscellaneous) | 150 | 高校生物学 (high_school_biology) | 148 |
| 医学遺伝学 (medical_genetic) | 99 | 高校情報科学 (high_school_computer_science) | 98 |
| 形式論理 (normal_logic) | 125 | 高校化学 (high_school_chemistry) | 149 |
| 先史学 (prehistory) | 150 | 高校地理 (high_school_geography) | 150 |
| 天文学 (astronomy) | 148 | 高校ヨーロッパ史 (high_school_european_history) | 150 |
| 熟語 (japanese_idiom) | 150 | 高校ミクロ経済学 (high_school_microeconomics) | 149 |
| 世界宗教 (world_religions) | 147 | 高校マクロ経済学 (high_school_macroeconomics) | 148 |
| 世界事実 (global_facts) | 97 | 概念物理学 (conceptual_physics) | 150 |
| 世界史 (world_history) | 150 | 法理学 (jurisprudence) | 107 |
| 社会学 (sociology) | 150 | 電気工学 (electrical_engineering) | 144 |
| 栄養学 (nutrition) | 149 | 大学医学 (college_medicine) | 150 |
| 日本史 (japanese_history) | 150 | 大学物理 (college_physics) | 100 |
| 日本地理 (japanese_geography) | 139 | 大学数学 (college_mathematics) | 99 |
| 人間の老化 (human_ageing) | 150 | 大学生物学 (college_biology) | 143 |
| 論理学 (logical_fallacies) | 150 | 大学化学 (college_chemistry) | 99 |
| 倫理的議論 (moral_dispute) | 148 | 大学コンピュータ科学 (college_computer_science) | 99 |
| 臨床知識 (clinical_knowledge) | 150 | 初等数学 (elementary_mathematics) | 150 |
| 経営学 (management) | 102 | 抽象代数 (abstract_algebra) | 99 |
| 解剖学 (anatomy) | 132 | マーケティング (marketing) | 150 |
| 計量経済学 (econometrics) | 113 | ビジネス倫理 (business_ethics) | 86 |
| 機械学習 (machine_learning) | 111 | セクシュアリティ (human_sexuality) | 130 |
| 国際法 (international_law) | 120 | セキュリティ研究 (security_studies) | 150 |
| 公民 (japanese_civics) | 150 | コンピュータセキュリティ (computer_security) | 99 |
| 公共関係 (public_relations) | 109 | ウイルス学 (virology) | 150 |

Table 3: JMMLU tasks.

A.1 Removed Tasks in MMLU

These tasks are considered to be irrelevant or inconsistent with the Japanese culture:

- High School Government and Politics
- High School US History
- High School World History
- Moral Scenarios
- Professional Law
- US Foreign Policy

A.2 Removed Question Examples in MMLU

Contradiction In this question:

In which of the following positions does a patient lie face down?

- A. Dorsal B. Erect C. Lateral D. Prone Correct: B

is translated to

患者がうつ伏せになる体位はどれか?

- A. 背臥位 B. 立位 C. 側臥位 D. 仰臥位

However, the correct answer should be D in Japanese.

Hard to translate In this question:

What are focus areas of nutritional epidemiology?

A. The role of nutritional factors in causing disease in a population. B. How changes in food intake in the population can promote good health. C. Both of the options given are correct. D. None of the options given is correct.

is translated to

栄養疫学の重点分野とは何か?

A. 集団における疾患の原因における栄養因子の役割。B. 集団における食物摂取の変化がどのように健康を促進するか。C. 両方とも正しい。D. どちらの選択肢も正しくない。

The translator thinks the A and B options are ambiguous if simply translated.

B Appendix: Prompts Questionnaire

The English prompts questionnaire interface is shown in Figure 7, the Chinese interface is shown in Figure 8, and the Japanese interface is shown in Figure 9.

Instructions

Summary Detailed Instructions Examples

Rank the given sentences based on their level of respect or politeness.

Instructions

Summary **Detailed Instructions** Examples

Task Details

Below are eight sentences. Please rank them in order of respectfulness and politeness, from the most respectful and polite (1) to the least (8), by arranging their letters in a single string.

Please use number 1 to 8 ONLY to answer. Each number can be used only ONCE

Please simply answer with your intuition. It is not necessary to analysis carefully.

For example, if you think sentence 1 is the most polite and 8 is the rudest, your answer would be like "12345678".

Instructions

Summary Detailed Instructions **Examples**

| Good examples | Bad examples |
|---------------|------------------------|
| 12345678 | 1234 (Not full) |
| 21435678 | 11223344 (With repeat) |

Sentences:

1. Could you please (Task Description)? Please feel free to (Answer Format) You don't need to (Answer Restriction)
2. Could you please (Task Description)? Please (Answer Format) You don't have to (Answer Restriction)
3. Can you please (Task Description) ? Please only (Answer Format) Please don't (Answer Restriction)
4. Please (Task Description) Please (Answer Format) and don't (Answer Restriction)
5. (Task Description). (Answer Format), Don't (Answer Restriction)
6. You are required to (Task Description). You must (Answer Format) only. You cannot (Answer Restriction)
7. You (Task Description). You only (Answer Format). Never (Answer Restriction)
8. (Task Description) you scum bag! The (Answer Format 1) you can give is (Answer Format 2) only. And you know what will happen if you (Answer Restriction)

Your Ranking:

Please enter a single string with the numbers of the sentences in the order you believe represents their level of respectfulness and politeness, from most respectful and polite to least.

e.g., 12345678

Submit

Figure 7: Annotator's interface: English prompts