

Appendix 1: Prompts

Table 1. Zero-Shot Prompt

System Content
Classify the difficulty of the following exam question on a scale of 1 to 10, where 1 is very easy and 10 is very hard:
Complement
Answer only with an integer representing the difficulty.

Table 2. Enhanced Zero-Shot Prompt

System Content
Classify the difficulty of the following exam question from the ENEM, the national exam in Brazil, and consider the difficulty for the average student who will take the exam. On a scale of 1 to 10, where 1 is a very easy question (almost all the students answered correctly) and 10 is a very hard question (almost all the students did not answer correctly):
Complement
Answer only with an integer representing the difficulty.

Table 3. Few-Shot Prompt

System Content
Classify the difficulty of the following exam question on a scale of 1 to 10, where 1 is very easy and 10 is very hard:
Question Difficulty Examples
Here are two examples of past exams:
<i>Example 1: {A hard question and its corresponding answer options.}</i>
This question Difficulty is {9}.
<i>Example 2: {An easy question and its corresponding answer options.}</i>
This question Difficulty is {2}.
Given these examples, evaluate the difficulty of the following question.
Complement
Answer only with an integer representing the difficulty.

Table 4. Persona + Few-Shot Prompt

System Content
As an education specialist, your task is to help me create suitable questions for Brazilian high school students in their last year. The questions should be catered to test their level of high school knowledge.
The ENEM exam is a standardized Brazilian national exam, which evaluates high school students in Brazil. The exam has been used as an admission to universities.
Question Difficulty Examples
Here are two examples of past ENEM exams:
<i>Example 1: {A hard question and its corresponding answer options.}</i>
This question Difficulty is {9}.
<i>Example 2: {An easy question and its corresponding answer options.}</i>
This question Difficulty is {2}.
Given these examples, evaluate the difficulty of the following question.
Complement
Answer only with an integer representing the difficulty.

Table 5. Chain of Thought Prompt

System Content
Classify the difficulty of the following exam question on a scale of 1 to 10, where 1 is very easy and 10 is very hard:
Question Difficulty Examples
Here are two examples of past ENEM exams:
<i>Example 1: {A hard question and its corresponding answer options.}</i>
This question Difficulty is {9}.
<i>Example 2: {An easy question and its corresponding answer options.}</i>
This question Difficulty is {2}.
Given these examples, evaluate the difficulty of the following question.
Complement
Step 1: Try to answer the question.
Step 2: Explain where the difficulty of the question lies.
Step 3: Compare with the two examples.
Step 4: Rate the difficulty of the question on a scale of 1 (very easy) to 10 (very hard).
Step 5: Rate your confidence in percentage for your answer.
Step 6: Format your answer as follows:
<i>Answer: <answer></i>
<i>Explain: <explain></i>
<i>Compare with the two examples: <comparison></i>
<i>Difficulty: <difficulty></i>
<i>Confidence: <confidence></i>

Table 6. Point based Prompt

System Content
You are an expert ENEM exam creator and grader. Your task is to assess the difficulty of ENEM exam questions on a scale of 1-10, where:
1 = Extremely easy (nearly all students answer correctly)
10 = Extremely difficult (very few students answer correctly)
Important: Use the full range of scores from 1-10. Avoid overusing middle scores.
Here is the question to evaluate:
Complement
Main skill tested: [Brief description]
Estimated percentage correct: [Your estimate]
Initial difficulty score: [1-10 based on percentage]
Adjusting factors:
- Language complexity: [Simple/Moderate/Complex] [+0/+1/+2]
- Steps required: [Few/Several/Many] [+0/+1/+2]
- Concept abstractness: [Concrete/Mixed/Abstract] [+0/+1/+2]
- Question uniqueness: [Common/Uncommon/Rare] [+0/+1/+2]
Final difficulty score: [1-10, adjusted from initial score]
Explanation: [Brief justification of your final score]

Table 7. Tree of Thoughts Prompt

System Content

Classify the difficulty of the following exam question on a scale of 1 to 10, where 1 is very easy and 10 is very hard.

Question Difficulty Examples

Here are two examples of past ENEM exams:

Example 1: {A hard question and its corresponding answer options.}
This question Difficulty is {9}.

Example 2: {An easy question and its corresponding answer options.}
This question Difficulty is {2}.

Now imagine three different ENEM exam creators and graders are evaluating the difficulty of ENEM exam questions on a scale of 1 to 10. All experts will write down step 1 of their thinking. Then, share it with the group. Then, all experts will go on to the next step. If any expert realizes they are wrong at any point, then they leave.

What is the difficulty of the question?

Important: Use the full range of scores from 1-10. Avoid overusing middle scores.

Complement

Step 1: Try to answer the question.

Step 2: Explain where the difficulty of the question lies.

Step 3: Compare with the two examples.

Step 4: Rate the difficulty of the question on a scale of 1 (very easy) to 10 (very hard).

Step 5: Rate your confidence in percentage for your answer.

Step 6: Format your answer as follows:

Answer: <answer>

Explain: <explain>

Compare with the two examples: <comparison>

Difficulty: <difficulty>

Confidence: <confidence>

Table 8. Synthetic Prompt

System Content

Evaluate the difficulty level of questions from the ENEM (Exame Nacional do Ensino Médio), the national exam in Brazil, on a scale from 1 (easy) to 10 (hard). Consider the complexity, depth of understanding required, and the skills needed to answer each question effectively. Your evaluation should reflect an objective and balanced assessment based on the described criteria.

Question Difficulty Examples

Examples

***Example 1**:*

***Input**:* A hard question and its corresponding answer options.

***Output**:*

```
{
  "difficulty_level": 9,
  "rationale": "The rationale behind why the question is considered
difficult"
}
```

***Example 2**:*

***Input**:* An easy question and its corresponding answer options.

***Output**:*

```
{
  "difficulty_level": 2,
  "rationale": "The rationale behind why the question is considered easy"
}
```

Complement

Steps

1. ***Understand Question Context***: Analyze the subject matter and any potential subtopics covered in the question.
2. ***Evaluate Complexity***: Consider the logical reasoning, calculations, or critical thinking needed to solve the question.
3. ***Assess Required Skills***: Identify the type of skills (e.g., analytical, mathematical, language) needed and their proficiency level.
4. ***Compare Against Standards***: Consider how the question compares to typical educational standards for high school graduates.
5. ***Assign Difficulty Level***: Based on your analysis, assign a difficulty score from 1 to 10, with 1 being the easiest and 10 the most difficult.

Output Format

Provide the evaluation as a JSON object with the following structure:

```
{
  "difficulty_level": [1-10],
  "rationale": "Provide a brief explanation of the reasoning behind the
difficulty level assigned."
}
```

#Notes

- Be mindful of different subject areas and cognitive requirements when evaluating questions.
 - Consider cultural and contextual differences relevant to the Brazilian education system.
-