### **Relevance:**

Please score the relevance of the response in relation to the context of the dialogue. Relevance refers to the degree of association between the response and the questions, topics, or information within the context.

Rating 1:

* Lacks information relevant to the dialogue content; the response has no direct connection with the context.
* The response might be vague, making it difficult to understand its relation to the dialogue.
* Insufficient detailed information is provided, making the connection between the response and the dialogue unclear.

Rating 2:

* Provides some information relevant to the dialogue content, but still lacks detailed connection.
* The response is somewhat relevant to the context but may require more specific information to clarify the connection.
* The degree of relevance in the response is average, not yet fully meeting the dialogue needs.
* The connection in the response is clearer than in lower scores, but there is room for improvement.

Rating 3:

* Provides clear, detailed, and comprehensive information that is directly relevant to the dialogue content.
* Uses specific facts, nouns, verbs, and time/place information from the context, making the relevance of the response to the dialogue very apparent.
* The connection between the response and the dialogue is very high, with no key details missed.
* The response is clear and concise, closely linked to the dialogue, fulfilling the dialogue needs.

### 

### 

### **Interestingness:**

Please score the interestingness of the response in relation to the context of the dialogue. The evaluation of interestingness usually involves the creativity, humor, engaging nature, or uniqueness of the reply.

Rating 1:

* The reply lacks interestingness, appearing monotonous and dull.
* It does not contain any creative or engaging elements.
* The reply fails to attract or arouse interest, lacking entertainment value.

Rating 2:

* The reply contains some interesting elements, but is still rather bland.
* There might be slight creativity or humor, but it is not sufficiently engaging.
* The reply arouses interest to some extent, but there is room for improvement.

Rating 3:

* The reply is very interesting, filled with creativity, humor, or engaging content.
* It includes engaging plots, humorous moments, or unique perspectives.
* The reply is pleasing and able to attract and maintain the audience's interest.
* An interesting reply typically makes the conversation more enjoyable and pleasant, enhancing the user experience.

### **Understandability:**

Please score the comprehensibility of the response in relation to the context of the dialogue. Comprehensibility refers to the ability of the reply to ensure that the recipient clearly understands the information conveyed.

Rating 1:

* The comprehensibility is extremely low; the content of the reply or text is difficult to understand.
* Uses vague, obscure, or highly technical language, making it hard for the average person to understand.
* Insufficient background information is provided, leaving readers unable to understand the context.

Rating 2:

* Comprehensibility is average; the content of the reply or text is understandable to a certain extent.
* Uses general language, but there may still be some vagueness or unclear expressions.
* Some background information is provided, but more details might be needed for complete understanding.

Rating 3:

* Comprehensibility is very high; the content of the reply or text is very easy to understand.
* Uses clear and concise language, avoiding obscure and difficult expressions.
* Provides sufficient background information, making the context easy to understand and satisfying the needs of the reader.

### **Overall Impression:**

Based on your impression of the quality of the response, rate the response on a scale of 1 to 3. When annotating, consider the content, clarity, relevance, and accuracy of the response. Assign a score that best reflects the overall quality of the dialogue in meeting these criteria.

Rating 1:

* The response is off-topic, inaccurate, or does not address the user's query.
* The language is unclear, confusing, or contains many grammatical errors.
* The response includes irrelevant or unnecessary information.

Rating 2:

* The response is somewhat relevant but may not fully address the user's query.
* The language is generally clear, but there may be minor grammatical errors or ambiguities.
* The response contains some relevant information but may miss key details or include minor inaccuracies.

Rating 3:

* The response accurately and completely addresses the user's query.
* The language is clear, concise, and free of grammatical errors.
* The response is well-structured, includes all relevant information, and demonstrates a thorough understanding of the topic.