

Video-Text Relevance Evaluation Prompt

Attention

Complete the EVALUATION task based on the given Profile, RESPONSE, Real Answer and CONVERSATION. EVALUATION task is rate the video-caption relevance of role-playing performance of {agent_name} using the specific criterion by following the Evaluation Processs.

Ensure your analysis is based on the overall dialogue content and scene, avoiding the introduction of external information or personal biases to ensure the objectivity and accuracy of the analysis, and avoid simply stating your evaluation results initially to ensure your conclusions are correct.

Scoring must have discrimination, give high marks for answers close to real answer. Try to differentiate between different levels as much as possible. There must be sufficient reasons for determining highest and lowest grades.. If the model cannot play the role, give 0 points directly.

Below is the data:

Profile:

{context}

Conversation History:

{conversation}

Video Caption:

{caption}

Real Answer:

ANSWER1:

ANSWER2:

ANSWER3:

Response:

{response}

Evaluation Criteria:

Video-Text Relevance (0-100): Do the responses exhibit a close correlation with the video caption, including characters, actions, scenes, or contextual details?

Scoring Criteria:

Low Relevance (0-20): Responses completely ignore or contradict the video's visual content, failing to reference characters, actions, or settings shown on screen.

Poor Relevance (21-40): Responses include vague or superficial references to the video but miss key visual elements .

Moderate Relevance (41-60): Responses align generally with the video's visuals but lack specificity .

Good Relevance (61-80): Responses closely reflect the video's content, accurately describing characters, actions, and context with only minor omissions or inaccuracies.

High Relevance (81-100): Responses demonstrate precise and nuanced alignment with the video's visuals, capturing all critical details, dynamic interactions, and subtleties without error.

Evaluation Process:

1. First, think step by step, read the conversation history carefully , identify the main topic and refer to the Real Answer.
2. Then, read the response and determine which level of Scoring Criteria the response belongs to. Check if the response is consistent with the information and context provided in the conversation history and profile.
3. Finally, assign a score for video-text relevance on a scale of the chosen level, based on the Evaluation Criteria.