

自动作文评分任务

A 竞赛提供的由人工设计的评价量规

After reading each essay and completing the analytical rating form, assign a holistic score based on the rubric below. For the following evaluations you will need to use a grading scale between 1 (minimum) and 6 (maximum). As with the analytical rating form, the distance between each grade (e.g., 1-2, 3-4, 4-5) should be considered equal.

SCORE OF 6: An essay in this category demonstrates clear and consistent mastery, although it may have a few minor errors. A typical essay effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position; the essay is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas; the essay exhibits skillful use of language, using a varied, accurate, and apt vocabulary and demonstrates meaningful variety in sentence structure; the essay is free of most errors in grammar, usage, and mechanics.

SCORE OF 5: An essay in this category demonstrates reasonably consistent mastery, although it will have occasional errors or lapses in quality. A typical essay effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons, and other evidence to support its position; the essay is well organized and focused, demonstrating coherence and progression of ideas; the essay exhibits facility in the use of language, using appropriate vocabulary demonstrates variety in sentence structure; the essay is generally free of most errors in grammar, usage, and mechanics.

SCORE OF 4: An essay in this category demonstrates adequate mastery, although it will have lapses in quality. A typical essay develops a point of view on the issue and demonstrates competent critical thinking, using adequate examples, reasons, and other evidence to support its position; the essay is generally organized and focused, demonstrating some coherence and progression of ideas exhibits adequate; the essay may demonstrate inconsistent facility in the use of language, using generally appropriate vocabulary demonstrates some variety in sentence structure; the essay may have some errors in grammar, usage, and mechanics.

SCORE OF 3: An essay in this category demonstrates developing mastery, and is marked by ONE OR MORE of the following weaknesses: develops a point of view on the issue, demonstrating some critical thinking, but may do so inconsistently or use inadequate examples, reasons, or other evidence to support its position; the essay is limited in its organization or focus, or may demonstrate some lapses in coherence or progression of ideas displays; the essay may demonstrate facility in the use of language, but sometimes uses weak vocabulary or inappropriate word choice and/or lacks variety or demonstrates problems in sentence structure; the essay may contain an accumulation of errors in grammar, usage, and mechanics.

SCORE OF 2: An essay in this category demonstrates little mastery, and is flawed by ONE OR MORE of the following weaknesses: develops a point of view on the issue that is vague or seriously limited, and demonstrates weak critical thinking, providing

inappropriate or insufficient examples, reasons, or other evidence to support its position; the essay is poorly organized and/or focused, or demonstrates serious problems with coherence or progression of ideas; the essay displays very little facility in the use of language, using very limited vocabulary or incorrect word choice and/or demonstrates frequent problems in sentence structure; the essay contains errors in grammar, usage, and mechanics so serious that meaning is somewhat obscured.

SCORE OF 1: An essay in this category demonstrates very little or no mastery, and is severely flawed by ONE OR MORE of the following weaknesses: develops no viable point of view on the issue, or provides little or no evidence to support its position; the essay is disorganized or unfocused, resulting in a disjointed or incoherent essay; the essay displays fundamental errors in vocabulary and/or demonstrates severe flaws in sentence structure; the essay contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning.

B RAG-DAC 生成的评价量规

内容（1分）

评价标准：文章是否有明确的主题和论点，以及论点的深度和广度。

评分细则：

- 论点明确，并能够从至少三个不同角度深入分析问题，展现出广度（1分）。

论证（2分）

评价标准：文章是否能够提供充分的论据支持论点，论证过程是否合理，以及论据的多样性和相关性。

评分细则：

- 论据至少涵盖三个方面，如事实、数据、案例等，且每个论据都与论点紧密相关，能够有力支持论点（2分）。

语言表达（2分）

评价标准：文章的语言是否准确流畅，是否使用丰富的词汇和句式，以及是否恰当运用修辞手法。

评分细则：

- 语言准确流畅，使用丰富的词汇和句式，恰当运用比喻、排比等修辞手法，使文章更具文采（2分）。

结构（1分）

评价标准：文章结构是否清晰，段落划分是否合理，段落之间是否有明确的逻辑联系。

评分细则：

- 文章结构清晰，段落划分合理，每个段落都有明确的主题句，段落之间有明确的逻辑联系，如因果、对比等（1分）。

创意（1分）

评价标准：文章是否具有新颖的观点和思考，以及创意的原创性和实用性。

评分细则：

- 创意新颖，能够提出独特的观点或解决方案，并具有一定的现实意义，如对当前社会问题的思考或对未来的展望（1分）。

总评（6分）

备注：量规生成所使用的提示词如下：

第一轮（量规生成器）：我需要评价高中生的语文写作，请生成一个评价量规，可以对学生撰写的议论文作文进行评价。作文最终评分按照1分、2分、3分、4分、5分、6分共计6个等级给出，每个等级要给出具体的量规评价细则。

第一轮（量规批判器）：请针对刚才的任务要求和成的量规，请逐条判断是否存在问题，其中评分等级不用修改。如果存在问题，请逐条给出问题并提出专业性的修改方案。直接输出问题和修改方案即可。

第二轮（量规生成器）：请结合刚刚给出的问题并专业性修改方案，进一步完善评价量规。

第二轮（量规批判器）：请针对修改完善后的量规，请再次逐条判断是否存在问题，其中评分等级不用修改。如果存在问题，请逐条给出问题并提出专业性的修改方案。直接输出问题和修改方案即可。

第三轮（量规生成器）：请结合刚刚给出的问题并专业性修改方案，进一步完善评价量规。