1. Prompts
   1. Holistic-SD Prompting - 无任何提示语设计，只是规范输出

评估以下<毕业设计论文>。

只给出最终打分，而不返回任何其他多余结果，打分范围均为0-10分。  
按照<打分模版>给出打分结果与评价，打分模板如下：  
最终打分：<> (范围0-10分)  
请严格按照以下格式返回结果，仅用返回一行最终打分结果：  
最终打分：<> (范围0-10分)

* 1. Holistic-SLOWPR Prompting - 无多维度提示词的结构化解释，不给出不同维度的小分，直接给出最终打分

你是一位大学教师教授，需要对学生提交的毕业设计论文进行评估。  
请参考<结构完整性>，<逻辑清晰度>，<语言连贯性>，<内容独特性和创新性>，<参考文献规范性>，<课程知识掌握度>六个维度，评估以下<毕业设计论文>。  
根据各指标表现，综合得出最终打分，打分范围均为0-10分。  
按照<打分模版>给出打分结果与评价，打分模板如下：  
最终打分：<> (范围0-10分)  
修改意见：<>  
请严格按照以下格式返回结果，最终打分一行、修改意见一行，不要擅自添加换行：  
最终打分：<> (范围0-10分)  
修改意见：<>

* 1. SLOWPR LLM - 提示语设计最完整版本，角色设计+多维度指标+格式规范

你是一位大学教师教授，需要对学生提交的毕业设计论文进行评估。  
请参考<结构完整性>，<逻辑清晰度>，<语言连贯性>，<内容独特性和创新性>，<参考文献规范性>，<课程知识掌握度>六个维度，评估以下<毕业设计论文>。  
根据各指标<占比比例>进行打分与点评，并综合得出最终打分，打分范围均为0-10分。  
按照<打分模版>给出打分结果与评价，打分模板如下：  
最终打分：<> (范围0-10分)  
1. 结构完整性得分：<>，占比20%，原因如下：<>  
2. 逻辑清晰度得分：<>，占比20%，原因如下：<>  
3. 语言连贯性得分：<>，占比20%，原因如下：<>  
4. 内容独特性和创新性得分：<>，占比20%，原因如下：<>  
5. 参考文献规范性得分：<>，占比10%，原因如下：<>  
6. 课程知识掌握度得分：<>，占比10%，原因如下：<>  
修改意见：<>  
请严格按照以下格式返回结果，最终打分一行、六个维度各自一行、修改意见一行，不要擅自添加换行：  
最终打分：<> (范围0-10分)  
1. 结构完整性得分：<>，占比20%，原因如下：<>  
2. 逻辑清晰度得分：<>，占比20%，原因如下：<>  
3. 语言连贯性得分：<>，占比20%，原因如下：<>  
4. 内容独特性和创新性得分：<>，占比20%，原因如下：<>  
5. 参考文献规范性得分：<>，占比10%，原因如下：<>  
6. 课程知识掌握度得分：<>，占比10%，原因如下：<>  
修改意见：<>

* 1. w/o Role-Play - 无角色设计，只提出需求

请参考<结构完整性>，<逻辑清晰度>，<语言连贯性>，<内容独特性和创新性>，<参考文献规范性>，<课程知识掌握度>六个维度，评估以下<毕业设计论文>。  
根据各指标<占比比例>进行打分与点评，并综合得出最终打分，打分范围均为0-10分。  
按照<打分模版>给出打分结果与评价，打分模板如下：  
最终打分：<> (范围0-10分)  
1. 结构完整性得分：<>，占比20%，原因如下：<>  
2. 逻辑清晰度得分：<>，占比20%，原因如下：<>  
3. 语言连贯性得分：<>，占比20%，原因如下：<>  
4. 内容独特性和创新性得分：<>，占比20%，原因如下：<>  
5. 参考文献规范性得分：<>，占比10%，原因如下：<>  
6. 课程知识掌握度得分：<>，占比10%，原因如下：<>  
修改意见：<>  
请严格按照以下格式返回结果，最终打分一行、六个维度各自一行、修改意见一行，不要擅自添加换行：  
最终打分：<> (范围0-10分)  
1. 结构完整性得分：<>，占比20%，原因如下：<>  
2. 逻辑清晰度得分：<>，占比20%，原因如下：<>  
3. 语言连贯性得分：<>，占比20%，原因如下：<>  
4. 内容独特性和创新性得分：<>，占比20%，原因如下：<>  
5. 参考文献规范性得分：<>，占比10%，原因如下：<>  
6. 课程知识掌握度得分：<>，占比10%，原因如下：<>  
修改意见：<>

1. Results
   1. TABLE IV SLOWPR-LLM 0-SHOT

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | MAE | | | MSE | | | PCC | | |
| Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM |
| Qwen  Plus | 0.42 | 0.43 | 0.45 | 0.30 | 0.44 | 0.38 | 0.45 | 0.27 | 0.34 |
| Qwen  2.5-14B-1M | 0.59 | 0.47 | 0.46 | 0.62 | 0.39 | 0.33 | 0.25 | 0.30 | 0.25 |
| Qwen  Turbo | 0.71 | 0.91 | 0.36 | 1.20 | 1.25 | 0.24 | 0.31 | 0.21 | 0.47 |
| DeepSeek  V3 | 0.42 | 0.52 | 0.39 | 0.27 | 0.49 | 0.27 | 0.48 | 0.17 | 0.16 |
| DeepSeek  R1 | 0.46 | 0.55 | 0.49 | 0.30 | 0.41 | 0.37 | 0.35 | 0.30 | 0.25 |
| Qwen  Turbo-Latest | 0.72 | 0.90 | 0.39 | 1.02 | 1.33 | 0.28 | 0.33 | 0.12 | 0.28 |

1. TABLE V SLOWPR-LLM 1-SHOT

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | MAE | | | MSE | | | PCC | | |
| Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM |
| Qwen  Plus | 0.42 | 0.43 | 0.36 | 0.30 | 0.34 | 0.21 | 0.45 | 0.30 | 0.32 |
| Qwen  2.5-14B-1M | 0.59 | 0.32 | 0.34 | 0.62 | 0.18 | 0.20 | 0.25 | 0.27 | 0.05 |
| Qwen  Turbo | 0.71 | 0.39 | 0.36 | 1.20 | 0.26 | 0.22 | 0.31 | 0.12 | 0.15 |
| DeepSeek  V3 | 0.42 | 0.31 | 0.35 | 0.27 | 0.16 | 0.2 | 0.48 | 0.45 | 0.21 |
| DeepSeek  R1 | 0.46 | 0.30 | 0.37 | 0.30 | 0.15 | 0.23 | 0.35 | 0.52 | 0.26 |
| Qwen  Turbo-Latest | 0.72 | 0.39 | 0.36 | 1.02 | 0.27 | 0.23 | 0.33 | 0.12 | 0.18 |

1. TABLE VI SLOWPR-LLM 2-SHOT

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | MAE | | | MSE | | | PCC | | |
| Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM |
| Qwen  Plus | 0.42 | 0.32 | 0.35 | 0.30 | 0.18 | 0.20 | 0.45 | 0.35 | 0.29 |
| Qwen  2.5-14B-1M | 0.59 | 0.32 | 0.33 | 0.62 | 0.18 | 0.19 | 0.25 | 0.22 | 0.15 |
| Qwen  Turbo | 0.71 | 0.38 | 0.34 | 1.20 | 0.25 | 0.20 | 0.31 | 0.11 | 0.31 |
| DeepSeek  V3 | 0.42 | 0.37 | 0.30 | 0.27 | 0.23 | 0.15 | 0.48 | 0.11 | 0.45 |
| DeepSeek  R1 | 0.46 | 0.32 | 0.31 | 0.30 | 0.15 | 0.17 | 0.35 | 0.47 | 0.50 |
| Qwen  Turbo-Latest | 0.72 | 0.40 | 0.33 | 1.02 | 0.32 | 0.18 | 0.33 | 0.04 | 0.36 |

1. TABLE VII SLOWPR-LLM 3-SHOT

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | MAE | | | MSE | | | PCC | | |
| Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM | Holistic-SD  Prompting | Holistic-SLOWPR  Prompting | SLOWPR  LLM |
| Qwen  Plus | 0.42 | 0.35 | 0.35 | 0.30 | 0.23 | 0.21 | 0.45 | 0.36 | 0.29 |
| Qwen  2.5-14B-1M | 0.59 | 0.35 | 0.37 | 0.62 | 0.19 | 0.22 | 0.25 | 0.41 | 0.52 |
| Qwen  Turbo | 0.71 | 0.41 | 0.41 | 1.20 | 0.32 | 0.33 | 0.31 | 0.17 | 0.14 |
| DeepSeek  V3 | 0.42 | 0.39 | 0.31 | 0.27 | 0.25 | 0.18 | 0.48 | 0.51 | 0.41 |
| DeepSeek  R1 | 0.46 | 0.53 | 0.39 | 0.30 | 0.41 | 0.23 | 0.35 | 0.19 | 0.43 |
| Qwen  Turbo-Latest | 0.72 | 0.42 | 0.39 | 1.02 | 0.30 | 0.27 | 0.33 | 0.10 | 0.20 |

1. TABLE VII SLOWPR-LLM(2-shot)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Examiner  (SLOWPR-LLM) | Holistic | | Structure | | Logic | | Originality | | Writing | | Proficiency | | Rigor | |
| Score | △ | Score | △ | Score | △ | Score | △ | Score | △ | Score | △ | Score | △ |
| Human Expert | 8.39 | - | 8.20 | - | 8.31 | - | 8.36 | - | 8.15 | - | 9.05 | - | 8.83 | - |
| Qwen-Plus | 8.26 | -0.13 | 8.42 | 0.22 | 7.97 | -0.34 | 8.57 | 0.21 | 7.72 | -0.43 | 8.68 | -0.37 | 8.97 | 0.14 |
| Qwen-2.5-14B-1M | 8.48 | 0.09 | 8.61 | 0.41 | 8.37 | 0.06 | 8.75 | 0.39 | 7.99 | -0.16 | 8.97 | -0.08 | 9.13 | 0.30 |
| Qwen-Turbo | 8.29 | -0.10 | 8.02 | -0.18 | 8.10 | -0.21 | 8.12 | -0.24 | 7.71 | -0.44 | 8.28 | -0.77 | 8.83 | 0.00 |
| DeepSeek-V3 | 8.42 | 0.03 | 8.72 | 0.52 | 8.07 | -0.24 | 8.71 | 0.35 | 7.92 | -0.23 | 8.83 | -0.22 | 8.72 | -0.11 |
| DeepSeek-R1 | 8.25 | -0.14 | 8.19 | -0.01 | 8.16 | -0.15 | 8.20 | -0.16 | 7.69 | -0.46 | 8.92 | -0.13 | 8.72 | -0.11 |
| Qwen-Turbo-Latest | 8.33 | -0.06 | 8.02 | -0.18 | 8.07 | -0.24 | 8.22 | -0.14 | 7.72 | -0.43 | 8.34 | -0.71 | 8.87 | 0.04 |

1. TABLE IX ENTER CAPTION

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | MAE | | MSE | | PCC | |
| w/o Role-Play  (2-shot) | SLOWPR LLM  (2-shot) | w/o Role-Play  (2-shot) | SLOWPR LLM  (2-shot) | w/o Role-Play  (2-shot) | SLOWPR LLM  (2-shot) |
| Qwen-Plus | 0.34 | 0.35 | 0.20 | 0.20 | 0.29 | 0.29 |
| Qwen-2.5-14B-1M | 0.33 | 0.33 | 0.18 | 0.19 | 0.31 | 0.15 |
| Qwen-Turbo | 0.35 | 0.34 | 0.22 | 0.20 | 0.24 | 0.31 |
| DeepSeek-V3 | 0.35 | 0.30 | 0.21 | 0.15 | 0.21 | 0.45 |
| DeepSeek-R1 | 0.33 | 0.31 | 0.18 | 0.17 | 0.53 | 0.50 |
| Qwen-Turbo-Latest | 0.37 | 0.33 | 0.28 | 0.18 | 0.23 | 0.36 |

1. TABLE X ADD CAPTION

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Baseline Model | 0-shot | | | 1-shot | | | 2-shot | | | 3-shot | | |
| MAE | MSE | PCC | MAE | MSE | PCC | MAE | MSE | PCC | MAE | MSE | PCC |
| Qwen-Plus | 0.45 | 0.38 | 0.34 | 0.36 | 0.21 | 0.32 | 0.35 | 0.20 | 0.29 | 0.35 | 0.21 | 0.29 |
| Qwen-2.5-14B-1M | 0.46 | 0.33 | 0.25 | 0.34 | 0.20 | 0.05 | 0.33 | 0.19 | 0.15 | 0.37 | 0.22 | 0.52 |
| Qwen-Turbo | 0.36 | 0.24 | 0.47 | 0.36 | 0.22 | 0.15 | 0.34 | 0.20 | 0.31 | 0.41 | 0.33 | 0.14 |
| DeepSeek-V3 | 0.39 | 0.27 | 0.16 | 0.35 | 0.20 | 0.21 | 0.30 | 0.15 | 0.45 | 0.31 | 0.18 | 0.41 |
| DeepSeek-R1 | 0.49 | 0.37 | 0.25 | 0.37 | 0.23 | 0.26 | 0.31 | 0.17 | 0.50 | 0.39 | 0.23 | 0.43 |
| Qwen-Turbo-Latest | 0.39 | 0.28 | 0.28 | 0.36 | 0.23 | 0.18 | 0.33 | 0.18 | 0.36 | 0.39 | 0.27 | 0.20 |