|  |  |  |  |
| --- | --- | --- | --- |
| 实验组别 | 数据集大小/组 | 单轮few-shot使用的样本数量/组 | 轮数epoch/轮 |
| A | 85 | 6 | 5 |
| B | 205 | 6 | 5 |
| C | 205 | 10 | 5 |
| D | 205 | 10 | 10 |

A # 85样本，已清洗 few-shot k=6 # 5 epoch

# 85样本，已清洗

# musicgen\_scene\_music\_pairs\_train\_v2.jsonl

# musicgen\_scene\_music\_pairs\_test\_v2.jsonl

# few-shot k=6

# 5epoch

📊 Round 1 Scores:

BLEU: 0.2699

ROUGE-L: 0.5816

METEOR: 0.6264

BERTScore\_F1: 0.8505

📊 Round 2 Scores:

BLEU: 0.2734

ROUGE-L: 0.5586

METEOR: 0.6458

BERTScore\_F1: 0.8774

📊 Round 3 Scores:

BLEU: 0.2475

ROUGE-L: 0.5473

METEOR: 0.6214

BERTScore\_F1: 0.8638

📊 Round 4 Scores:

BLEU: 0.2924

ROUGE-L: 0.5898

METEOR: 0.6470

BERTScore\_F1: 0.8673

📊 Round 5 Scores:

BLEU: 0.2611

ROUGE-L: 0.5661

METEOR: 0.6340

BERTScore\_F1: 0.8533

B # 205样本，已清洗 few-shot k=6 # 5 epoch

# 205样本，已清洗

# musicgen\_scene\_music\_pairs\_all\_cleaned\_train.jsonl

# musicgen\_scene\_music\_pairs\_all\_cleaned\_test.jsonl

# few-shot k=6

# 5 epoch

📊 Round 1 Scores:

BLEU: 0.2418

ROUGE-L: 0.5160

METEOR: 0.5579

BERTScore\_F1: 0.8600

📊 Round 2 Scores:

BLEU: 0.2402

ROUGE-L: 0.5478

METEOR: 0.6139

BERTScore\_F1: 0.8638

📊 Round 3 Scores:

BLEU: 0.2441

ROUGE-L: 0.5429

METEOR: 0.6083

BERTScore\_F1: 0.8667

📊 Round 4 Scores:

BLEU: 0.2582

ROUGE-L: 0.5607

METEOR: 0.6200

BERTScore\_F1: 0.8645

📊 Round 5 Scores:

BLEU: 0.2679

ROUGE-L: 0.5618

METEOR: 0.6431

BERTScore\_F1: 0.8684

C # 205样本，已清洗 few-shot k=10 # 5 epoch

# 205样本，已清洗

# musicgen\_scene\_music\_pairs\_all\_cleaned\_train.jsonl

# musicgen\_scene\_music\_pairs\_all\_cleaned\_test.jsonl

# few-shot k=10

# 5 epoch

📊 Round 1 Scores:

BLEU: 0.2732

ROUGE-L: 0.5741

METEOR: 0.6377

BERTScore\_F1: 0.8676

📊 Round 2 Scores:

BLEU: 0.2735

ROUGE-L: 0.5699

METEOR: 0.6284

BERTScore\_F1: 0.8686

📊 Round 3 Scores:

BLEU: 0.2844

ROUGE-L: 0.5780

METEOR: 0.6379

BERTScore\_F1: 0.8658

📊 Round 4 Scores:

BLEU: 0.2772

ROUGE-L: 0.5814

METEOR: 0.6403

BERTScore\_F1: 0.8680

📊 Round 5 Scores:

BLEU: 0.2764

ROUGE-L: 0.5778

METEOR: 0.6258

BERTScore\_F1: 0.8664

D # 205样本，已清洗 few-shot k=10 # 10 epoch

# 205样本，已清洗

# musicgen\_scene\_music\_pairs\_all\_cleaned\_train.jsonl

# musicgen\_scene\_music\_pairs\_all\_cleaned\_test.jsonl

# few-shot k=10

# 10 epoch

📊 Round 1 Scores:

BLEU: 0.2646

ROUGE-L: 0.5655

METEOR: 0.6516

BERTScore\_F1: 0.8623

📊 Round 2 Scores:

BLEU: 0.2927

ROUGE-L: 0.5851

METEOR: 0.6373

BERTScore\_F1: 0.8706

📊 Round 3 Scores:

BLEU: 0.2575

ROUGE-L: 0.5571

METEOR: 0.6107

BERTScore\_F1: 0.8634

📊 Round 4 Scores:

BLEU: 0.2708

ROUGE-L: 0.5646

METEOR: 0.6279

BERTScore\_F1: 0.8621

📊 Round 5 Scores:

BLEU: 0.2746

ROUGE-L: 0.5791

METEOR: 0.6328

BERTScore\_F1: 0.8614

📊 Round 6 Scores:

BLEU: 0.2824

ROUGE-L: 0.5835

METEOR: 0.6540

BERTScore\_F1: 0.8643

📊 Round 7 Scores:

BLEU: 0.2751

ROUGE-L: 0.5769

METEOR: 0.6422

BERTScore\_F1: 0.8750

📊 Round 8 Scores:

BLEU: 0.2757

ROUGE-L: 0.5730

METEOR: 0.6349

BERTScore\_F1: 0.8626

📊 Round 9 Scores:

BLEU: 0.2637

ROUGE-L: 0.5532

METEOR: 0.6356

BERTScore\_F1: 0.8626

📊 Round 10 Scores:

BLEU: 0.2749

ROUGE-L: 0.5704

METEOR: 0.6362

BERTScore\_F1: 0.8671

|  |  |
| --- | --- |
| 实验组别 | 描述 |
| C（我的模型） | 基于Qwen2.5-14B模型，使用few-shot、多轮对话提示生成 |
| E（zero-shot） | 基于Qwen2.5-14B模型，不使用few-shot，直接生成 |
| F（Qwen2.5-7B） | 基于Qwen2.5-7B模型，使用few-shot、多轮对话提示生成 |

E # 205样本，已清洗 few-shot k=10 # 5 epoch zero-shot

# 205样本，已清洗

# musicgen\_scene\_music\_pairs\_all\_cleaned\_train.jsonl

# musicgen\_scene\_music\_pairs\_all\_cleaned\_test.jsonl

# few-shot k=10

# 5 epoch

📊 Round 1 Scores:

BLEU: 0.0038

ROUGE-L: 0.0842

METEOR: 0.0836

BERTScore\_F1: 0.7496

📊 Round 2 Scores:

BLEU: 0.0029

ROUGE-L: 0.0846

METEOR: 0.0722

BERTScore\_F1: 0.7404

📊 Round 3 Scores:

BLEU: 0.0040

ROUGE-L: 0.0899

METEOR: 0.0820

BERTScore\_F1: 0.7480

📊 Round 4 Scores:

BLEU: 0.0040

ROUGE-L: 0.0812

METEOR: 0.0803

BERTScore\_F1: 0.7485

📊 Round 5 Scores:

BLEU: 0.0034

ROUGE-L: 0.0870

METEOR: 0.0760

BERTScore\_F1: 0.7427