

20260918-????????

0. ??????????

LLM

- : 60KB 1.2
- 1 25KB 2 54KB 2 28KB 3 36KB 1 3
- 1 129KB 1 102KB 1 NVFP4 12
- 8bit: 5,723 278 36,272 2,062

: ① ②
③1 NG Haiku Luna

1. ?????

6 Opus
Opus
NG 3 3 NG

- 1: codex_control_lab f7c3862 2 25KB
- 2: mox_system f3e69844 hub.py contract.json 54KB
- 1: alexa_cfw00 6e7e0f0 2 129KB
- 2: codex_control_lab e3cb8a4 1 28KB
- 3: mox_system 6a3aa19c 222 36KB
- 1: mox_system 0e1e43b6 7 102KB

Qwen3.8 NVFP4 gpt-oss:120b Qwen3.8 8bit
Claude Haiku

2. ?????

XXXXXXXXXXXXXXXXXXXX

項目	1	2	1	2	3	1
NVFP4 項目	2項目	項目	項目	3項目	2項目	項目
NVFP4 項目	2項目	項目	項目	項目	項目	項目
gpt-oss 項目	2項目	項目	項目	項目	項目	項目
gpt-oss 項目	1項目	項目	項目	項目	項目	項目
8bit 項目	1項目	1項目	項目	項目 項目	項目	項目
Claude Haiku 項目	1項目	1項目	項目	項目 1項目 NG 項目	項目 1項目 NG 項目	3項目

項目 22項目 11項目 9項目 2項目

3. ???????

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Haiku

項目	項目	項目	項目	項目
NVFP4 項目	1	2	152.4	11,438
NVFP4 項目	2	3	498.1	33,355
NVFP4 項目	1	3	2,358.0	109,021
NVFP4 項目	2	3	291.6	20,415
NVFP4 項目	3	2	137.9	9,214
NVFP4 項目	1	3	1,048.9	61,062
NVFP4 項目	1	2	189.1	13,906
NVFP4 項目	2	3	870.8	52,491
gpt-oss 項目	1	2	176.7	11,869
gpt-oss 項目	2	3	539.1	33,348
gpt-oss 項目	1	1	92.1	6,251
gpt-oss 項目	2	3	517.6	32,256
8bit 項目	1	1	278.1	5,723
8bit 項目	2	1	556.2	11,124

-  : NVFP4  1  353/353  60 
 NVFP4  1  64/64  2  NVFP4  2  110/110  30 
-  NG  : gpt-oss  2 1  110/110  1

-  NG


[illegible]

7. ???????????

項目	単位	単位換算
1GB	1024MB	1024MB
1MB	1024KB	1024KB
1KB	1024B	1024B

-  1  NG  Haiku  Luna  1  3
-  30 
- 
 1  7  95 
- 


100






1 
2 
2 













































































































































- : 8bit 1 2 2026-09-18
 8bit 8bit

- : NVFP4 1 1 gpt-oss 1 1

- : Claude Haiku 2 3 Qwen3.8 bf16

- [] : workflow/records/llm_coding_eval_20260917/[] prompts[] answers
[] checks[] reviews[] progress.md
- [] : workflow/records/llm_coding_eval_20260917/review_notes.md
- [] : harness/coding_runner.py[] harness/answer_diff.py
[] harness/run_q8_coding.py 8bit []
- [] : workflow/plans/plan_llm_coding_eval_260917.md

☐☐☐☐ #3
☐ : 2026-09-18 02:23:42 UTC (TRANS-☐)
☐ : 2026-09-18 02:27:44 UTC (TRANS-☐)