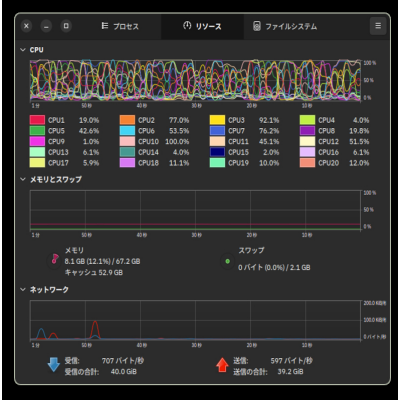




Llama3 70b

Description

Ubuntu 22.04 LTS

PC CPU Core i5 13500 RAM 64GB GPU RTX A4000 OS Ubuntu 22.04 LTS

AI RAG Kubernetes GPU LLM

CPU 30% AI 70b RAM 64GB



GPUã•16GBã®VRAMã??ã½¿ã•?ã??ã??æ??ã•?ã•§ã•?ã??æ¶?è²»é?»å??ã•æ??ã•§ã•§50Wã»ã•©ã•
§ã•?ã•®ã•§ã•??ã?»ã??è»¢ã•ã•?ã•?ã•?ã•?ã•§ã•ã•?ã•??ã•¼ã•?ã•??ã•??

```

x _ □ [F1] ktsubaki@ZIKUU-PC-02: ~
Every 1.0s: nvidia-smi ZIKUU-PC-02: Thu Jun 6 06:47:24 2024

+-----+
| NVIDIA-SMI 545.29.06                  Driver Version: 545.29.06      CUDA Version: 12.3      |
+-----+-----+
| GPU  Name                Persistence-M | Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |      Memory-Usage | GPU-Util  Compute M. |
|                                           | MIG M.         |
+-----+-----+
|    0  NVIDIA RTX A4000             Off | 00000000:01:00.0  On |          Off         |
| 41%    57C    P2              48W / 140W | 15341MiB / 16376MiB |      1%      Default |
|                                           |                  N/A |
+-----+-----+

+-----+
| Processes: |
| GPU   GI    CI          PID    Type    Process name                        GPU Memory |
|   ID   ID           |                     |              Usage      |
+-----+-----+
|    0   N/A   N/A       8963     G      /usr/lib/xorg/Xorg                  130MiB |
|    0   N/A   N/A       9093    C+G    ...libexec/gnome-remote-desktop-daemon 161MiB |
|    0   N/A   N/A       9142     G      /usr/bin/gnome-shell                259MiB |
|    0   N/A   N/A      10754     G      ...54,262144 --variations-seed-version 130MiB |
|    0   N/A   N/A      14202     G      gnome-control-center                 3MiB |
|    0   N/A   N/A      32009     C      ...unners/cuda_v11/ollama_llama_server 14634MiB |
+-----+

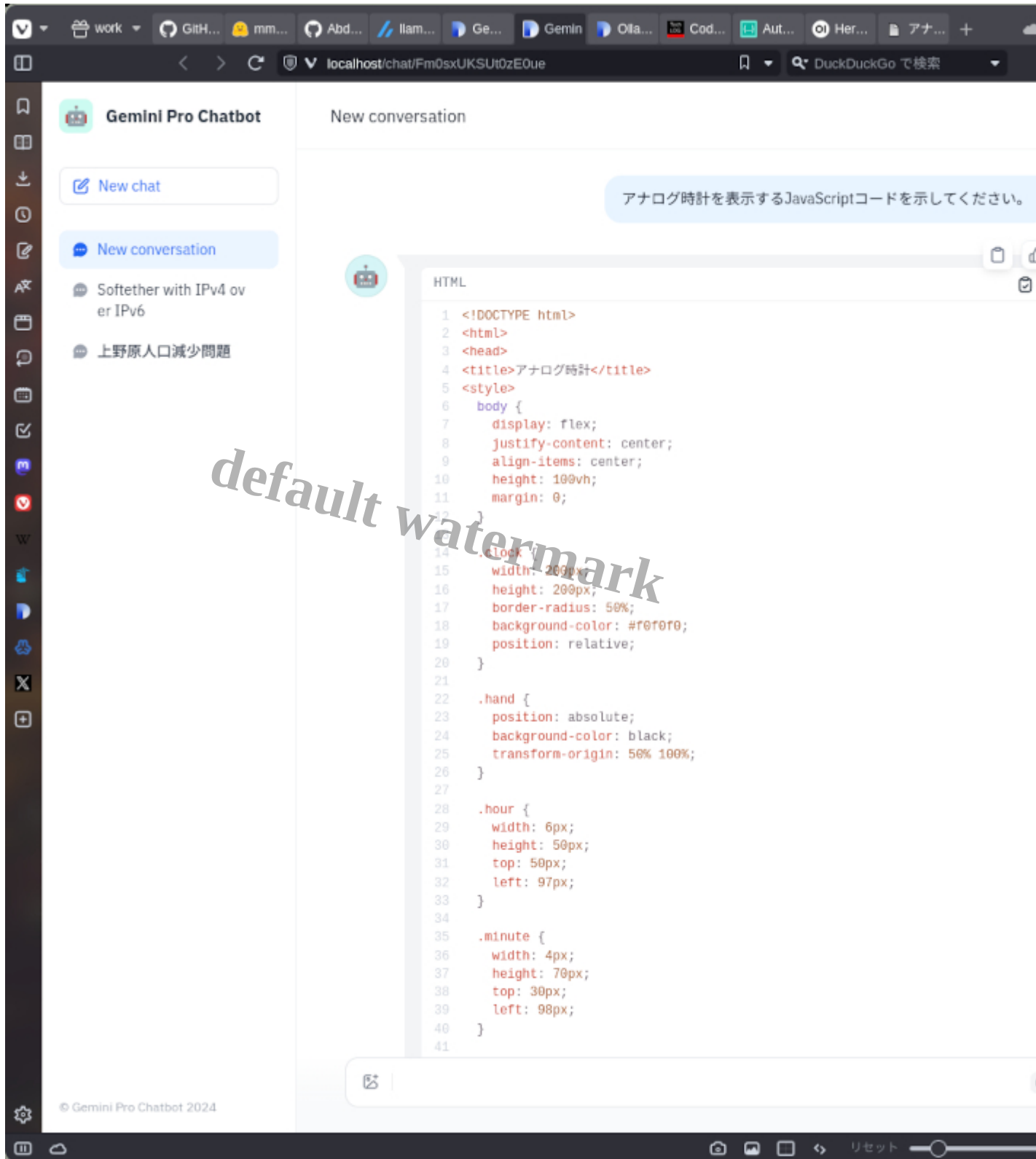
```

A4000ã¯æ¶¸?è²»é?«å??ã·æ?§ã·ç?®ã·§è?·ä¹ æ?§ã·?ã·?ã??ã·?ã·?ã·ª·?ã·ªã·®ã·?é å??ã·§ã·?ã·?ã·?ã·?ã·?ã·?
 »ã·©é??ã·?GPUã·§ã·-ã·ã·?ã·®ã·§ã·?ã·?ã·£ã·é??ã·?GPUã·é??ã·?CPUã·®çµ?ã·¿ã·?ã·?ã·?ã·ªã·?ã·
 ã·?ã·¶½¿ã·?ã·?æ??ã·?ã·«ã·ªã·?ã·®ã·§ã·-ã·ã·?ã·§ã·?ã·?ã·?ã·?ã·?ã·?ã·?ã·?ã·8bã· ã·
 ¨ã·¹ã·?ã·?-ã·¹ã·?ã·¾¼ã·£ã·?ã·?ã·ªã·?é«é?ã·ªã·?-ã·¹ã·?ã·³ã·¹ã·?ã·¾¼ã·?ã·?ã·?ã·|ã·?ã·?ã·®ã·¾¼ã·¾¼ã·
 §ã·?ã·½¿ã·?ã·?æ??ã·?ã·§ã·?ã·?

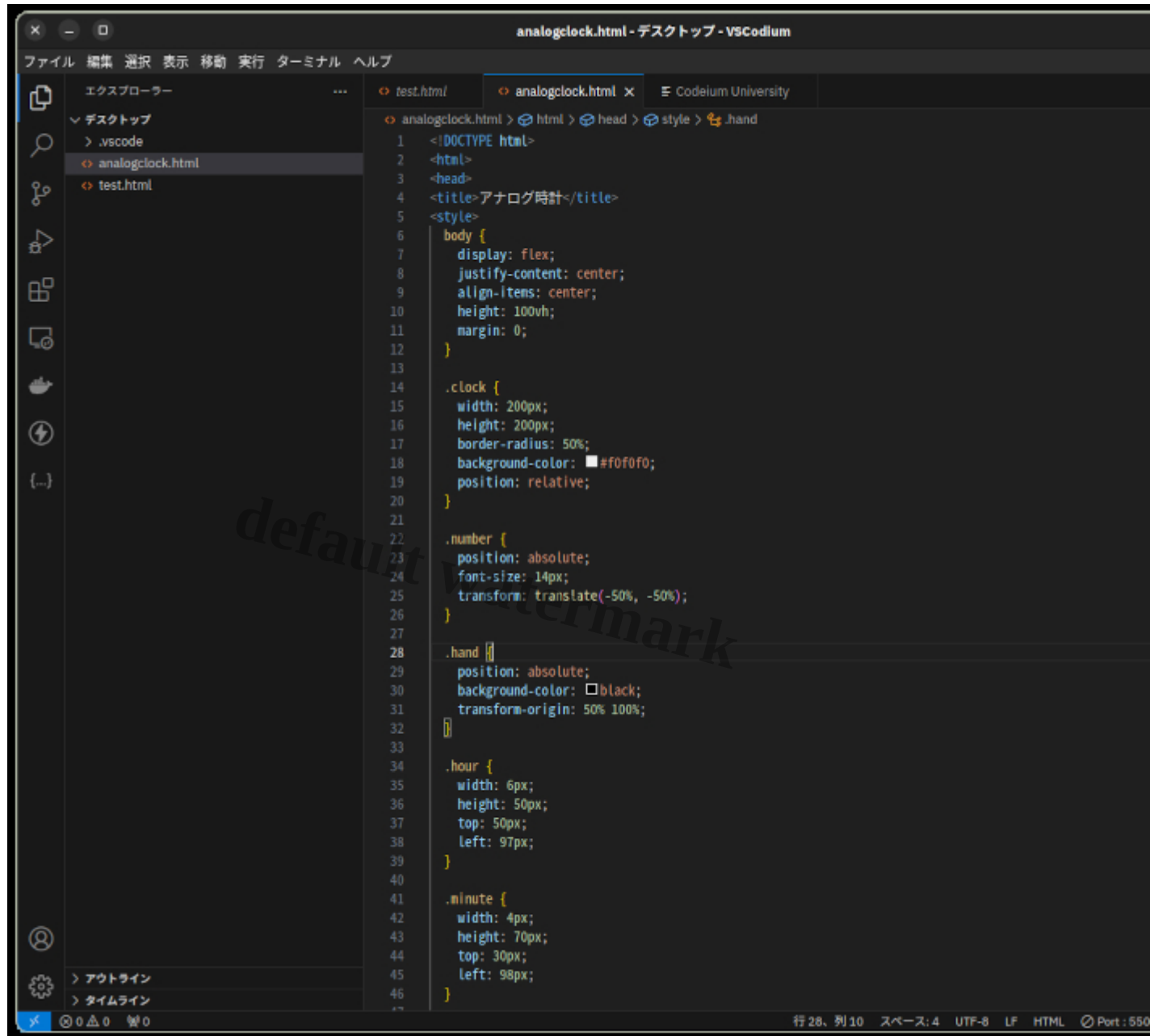
ã•jã•ªã•¿ã•«ã•?ã•??ã•-ã•??ã•\$ã•??ã•ã•°æ•??è•"?ã•??èj•çº°ã•?ã•??JavaScriptã•?³ã•?¼ã•??ã•??çº°ã•?ã•!ã••ã•
ã•?ã•?ã•??ã•?ã••ã•?ã•?è³ªã•?ã•??ã•?ã•?ã••ã••ã•®ã•??ã•®ã•§ã•?ã•??

å•?ã•?ã•?ã•?ã•?8bă•Să??è©|ă•?ă•¾ă•?ă•?ă•?ă•?ă•?ă•?ă•¼ă•?ă•?ă•!ă•?ă•?ă•?ă•?ă•?ă•Sç??æ?
 ä•?ä??ä•?ä³ä½ä⁻ä¾ä•£ä•?ä•?ä•?ä•?ä•?ä•?ä•?

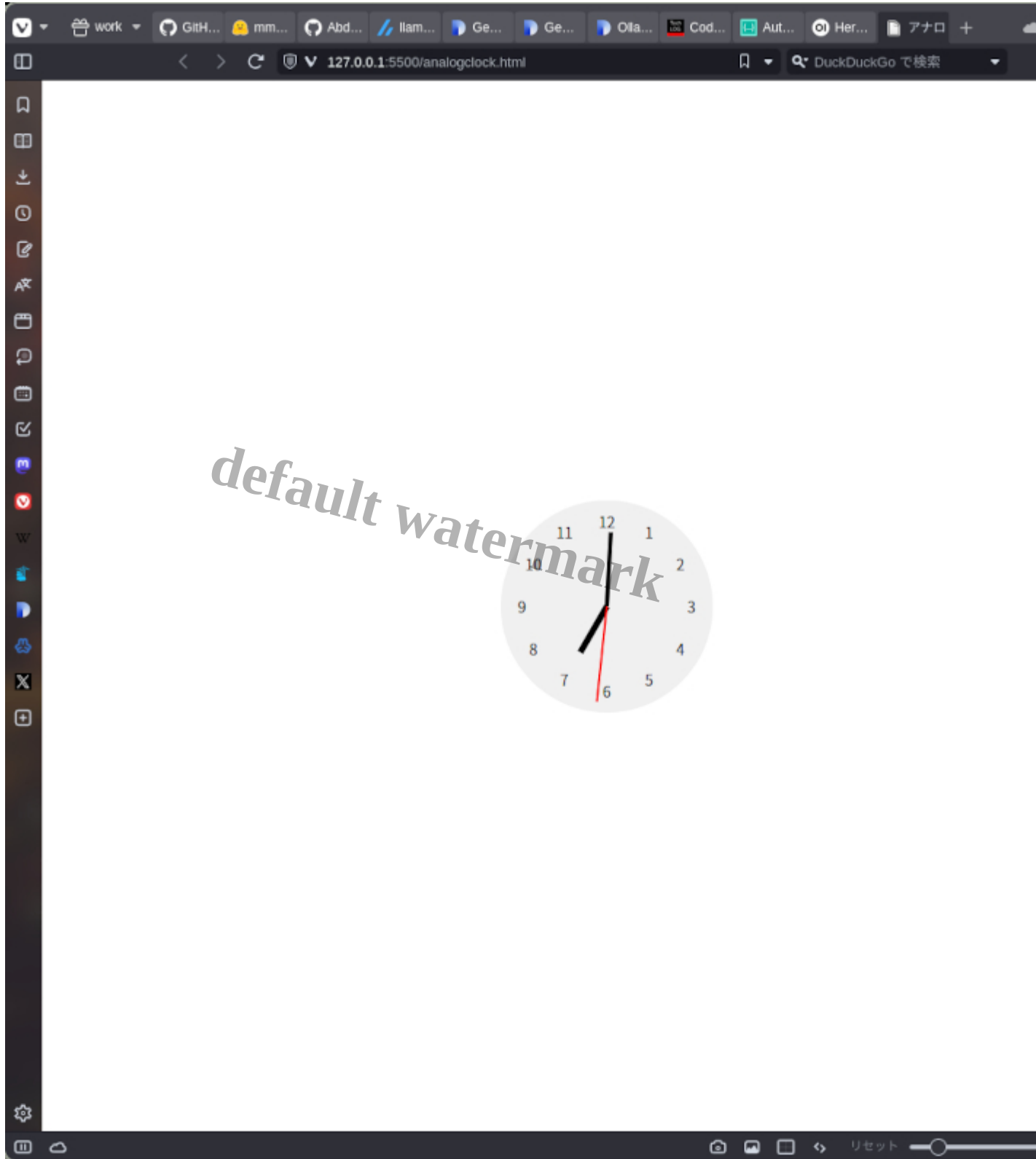
[illegible]



ç??æ?•ã?•ã??ã?•ã?¼ã?¼ã?¼ã?¾ã?¾ã?¾VS Codiumã«•ã³ã??ã??ã?•ã?|ä¸•å-
?ã??ã?¼½ã?¼¼ã?¼ã³ã?¼¼ã???ã«æ??ã·â ¥ã??ã|ã?•ã¾ã?ã??ã??



ã•ã•?ã•|ã•@?èj?ã??æ£ã•?ã•ã•?£ã•??ã•ã•°æ??è""ã•?èj"ç°ã•?ã•??ã•|ã•??ã•ã•¾ã•?ã•??



Gemini 1.5 Proã•®å´â•?ã¬HTML+CSS+JavaScriptã?ä½²ã•ã•ã£ã•?ã³ã¹¼ã??ã???ç??æ?•ã?ã|ã
•ã?ã•Llama3ã•ç??æ?•ã?ã•ã??ã•®ã??ã??èªå??ã§ã•ã?ã??

© Gemini - All Rights Reserved.

Llama3ã•?ç??æ?•ã•?ã•?ã³ã²¼ã??ã«ã¯CSSã?å«ã³¾ã??ã|ã•?ã¾ã•?ã??ãSã•?ã•?ã•?ã•
ã??CSSã??å«ã•?ã•|ã•ã•ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?
«æ??å?ã«CSSã??å«ã?ã•|æ¬²ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã•?ã£ã¯CSSç?jã•?ã•

© 2024 Gemini 1.5 Pro. All rights reserved. Gemini 1.5 Pro is a trademark of Google LLC. Gemini 1.5 Pro is a large language model trained by Google.

Category

- 1. AI
- 2. Dify

Tags

- 1. AI
- 2. Dify
- 3. Llama3
- 4. LLM
- 5. Ollama
- 6. Gemini 1.5 Pro
- 7. Gemini 1.5 Pro

Date Created

2024-06-06

Author

kazuo-tsubaki

default watermark