

llama.cpp को आजमा रहे हैं

जब आप llama.cpp को किसी मॉडल के साथ चलाने का प्रयास करते हैं, तो आपको इस तरह की त्रुटि का सामना करना पड़ सकता है:

```
(py311) lzwjava@Zhiweis-MacBook-Air llama.cpp % ./main -m models/7B/Phi-3-mini-4k-instruct-q4.gguf
main: build = 964 (f3c3b4b)
main: seed = 1737736417
llama.cpp: loading model from models/7B/Phi-3-mini-4k-instruct-q4.gguf
error loading model: unknown (magic, version) combination: 46554747, 00000003; is this really a GGML file?
llama_load_model_from_file: failed to load model
llama_init_from_gpt_params: error: failed to load model 'models/7B/Phi-3-mini-4k-instruct-q4.gguf'
main: error: unable to load model
```

यह त्रुटि आमतौर पर llama.cpp की स्थापना या मॉडल फ़ाइल में किसी समस्या का संकेत देती है।

एक सामान्य समाधान यह है कि llama.cpp को Homebrew का उपयोग करके स्थापित करें:

```
brew install llama.cpp
```

यह सुनिश्चित करता है कि आपके पास लाइब्रेरी का एक संगत संस्करण है।

यहां कुछ उपयोगी संसाधन दिए गए हैं:

- llama.cpp GitHub
- llama.cpp Homebrew
- llama.cpp Docker
- llama.cpp Windows
- llama.cpp Linux