Powered by backend.AI

PALI

Performant **A**I
Launcher for **I**nference

Download the model with a single click

One more click to launch your own service

Struggling with the setup to run your AI models?

Public institutions / Internal service operators
“Need easy access to models within organization users.”

Startups / Service planners
“Generative AI APIs are too pricey and complicated."

Developers / Researchers / AI beginners
“GPU Setup? Docker? Configuration is way too complex."

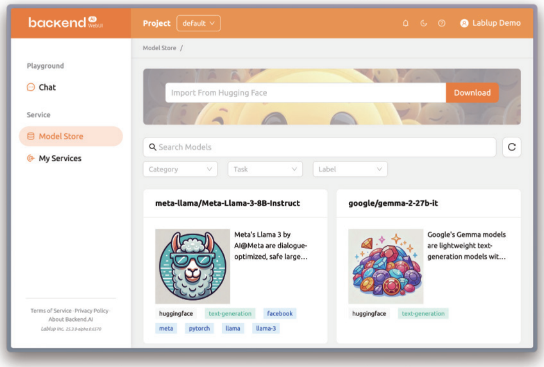
Start your AI journey with PALI in just one click.

lablup

Model Store

Various models in one place

- Provides curated catalog of public GenAI models
- Supports importing models from Hugging Face
- Supports NVIDIA NIM™ Inference Microservices
- Allows uploading user custom models
- Chat UI for instant model testing



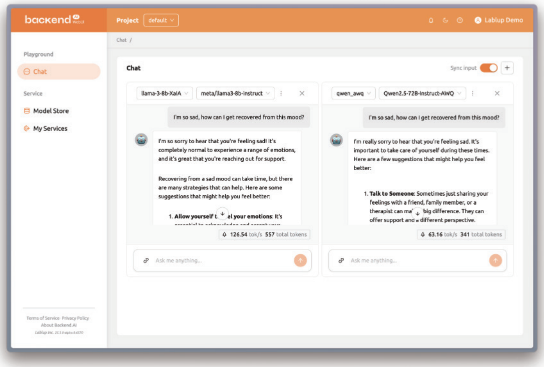
Meta's Llama 3 by AI@Meta are dialogue-optimized, safe large...

Google's Gemma models are lightweight text-generation models wit...

Playground

Instantly run selected models

- Test language models tailored to your service
- Fuse: combine multiple outputs for the best answer
- Compare models side by side with Multi-chat UI
- Supports image-based multimodal input/output
- Create, save, and reuse custom prompts



I'm so sad, how can I get recovered from this mood?

I'm so sorry to hear that you're feeling sad! It's completely normal to experience a range of emotions, and it's great that you're reaching out for support.

Recovering from a sad mood can take time, but there are many strategies that can help. Here are some suggestions that might help you feel better:

1. **Allow yourself t**[illegible]**el your emotions**: It's

I'm so sad, how can I get recovered from this mood?

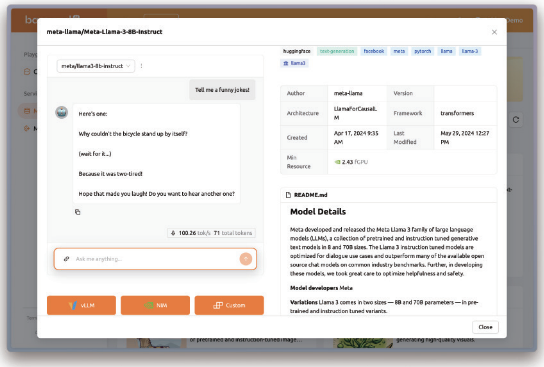
I'm really sorry to hear that you're feeling sad. It's important to take care of yourself during these times. Here are a few suggestions that might help you feel better:

1. **Talk to Someone**: Sometimes just sharing your feelings with a friend, family member, or a therapist can ma[illegible] big difference. They can offer support and a different perspective.

Model Player

Instantly run selected models

- Click & Play: Fast model execution
- Supports inference runtimes like vLLM / NIM / Hugging Face TGI
- Provides endpoints to integrate models with external services



Tell me a funny jokes!

Here's one:

Why couldn't the bicycle stand up by itself?

(wait for it...)

Because it was two-tired!

Hope that made you laugh! Do you want to hear another one?

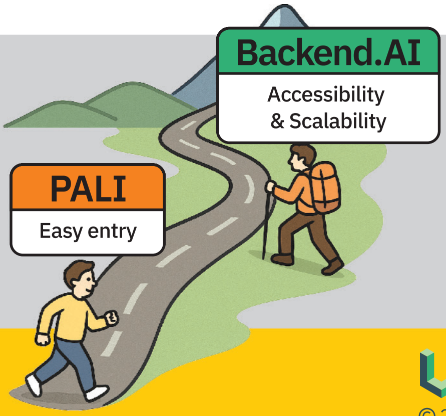
Author	meta-llama	Version	
Architecture	LlamaForCausalLM	Framework	transformers
Created	Apr 17, 2024 9:35 AM	Last Modified	May 29, 2024 12:27 PM
Min Resource	2.43 fGPU		

Model Details

Meta developed and released the Meta Llama 3 family of large language models (LLMs), a collection of pretrained and instruction tuned generative text models in 8 and 70B sizes. The Llama 3 instruction tuned models are optimized for dialogue use cases and outperform many of the available open source chat models on common industry benchmarks. Further, in developing these models, we took great care to optimize helpfulness and safety.

Model developers Meta

Variations Llama 3 comes in two sizes — 8B and 70B parameters — in pre-trained and instruction tuned variants.



PALI is your starting point for GenAI — launch complex AI setups in just two clicks. Works seamlessly across on both on-prem and cloud environments.

New to AI? Start simple with PALI.
Power user? Go deeper with Backend.AI.

Start your AI journey with Lablup.

lablup

contact@lablup.com
https://www.backend.ai
https://github.com/Lablup/backend.ai