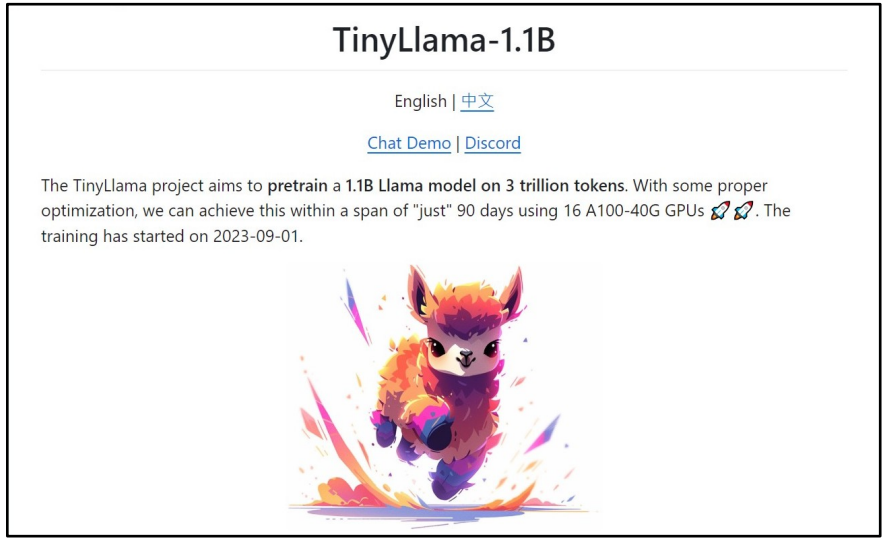
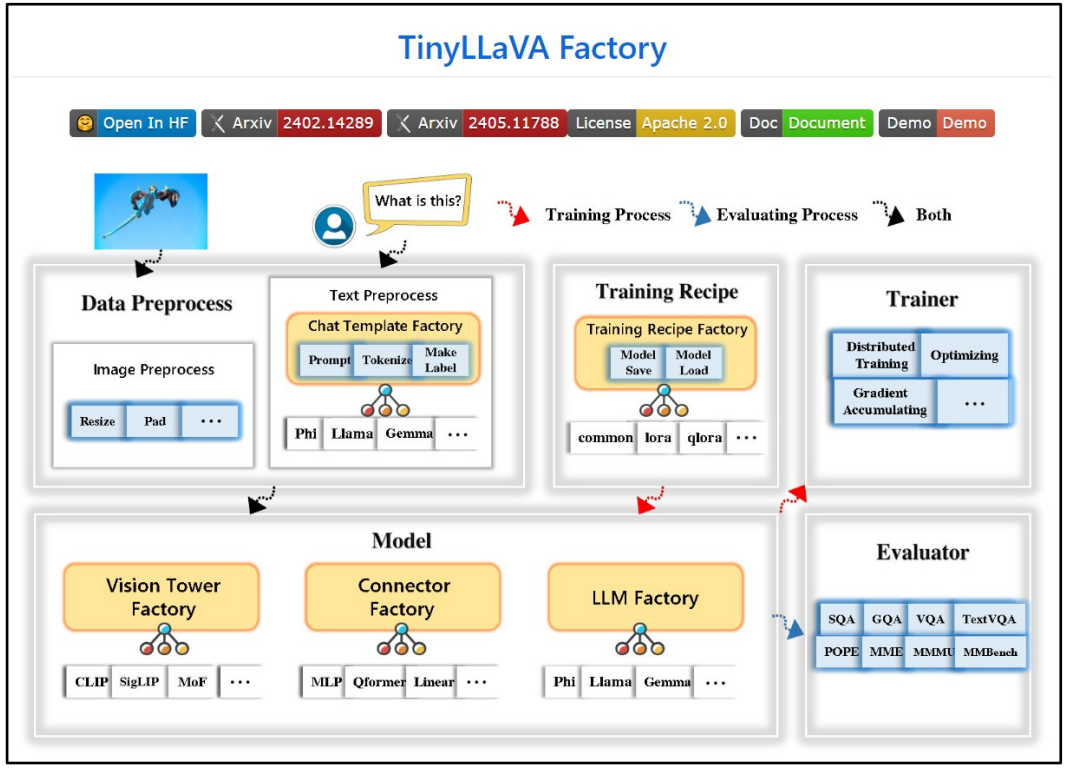
TinyLLaMA, TinyLLaVA, TinyLLaVA Factory

余方國 博士
fangkuoyu@gmail.com

TinyLLaMA, TinyLLaVA, TinyLLaVA Factory



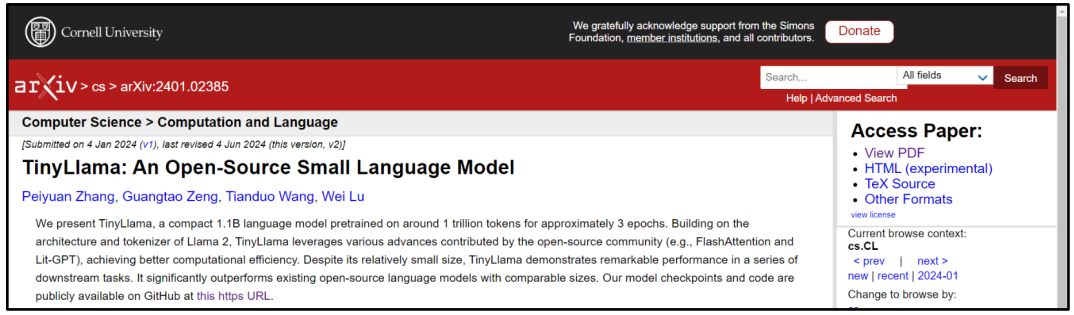
LLM

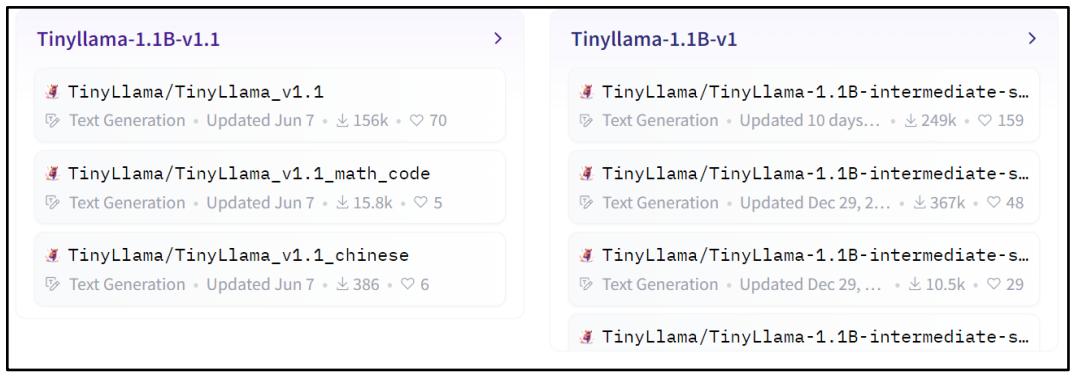


VLMs

https://github.com/jzhang38/TinyLlama
https://github.com/TinyLLaVA/TinyLLaVA_Factory

TinyLLaMA





Overview

- https://arxiv.org/abs/2401.02385
- https://github.com/jzhang38/TinyLlama
- https://huggingface.co/TinyLlama

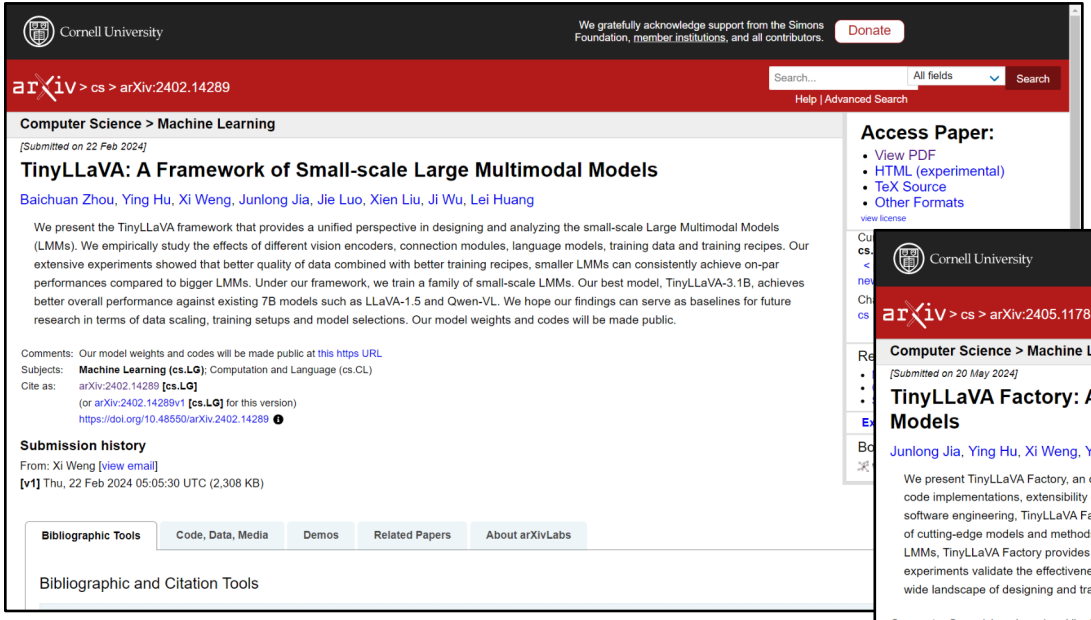
TRL SFTTrainer

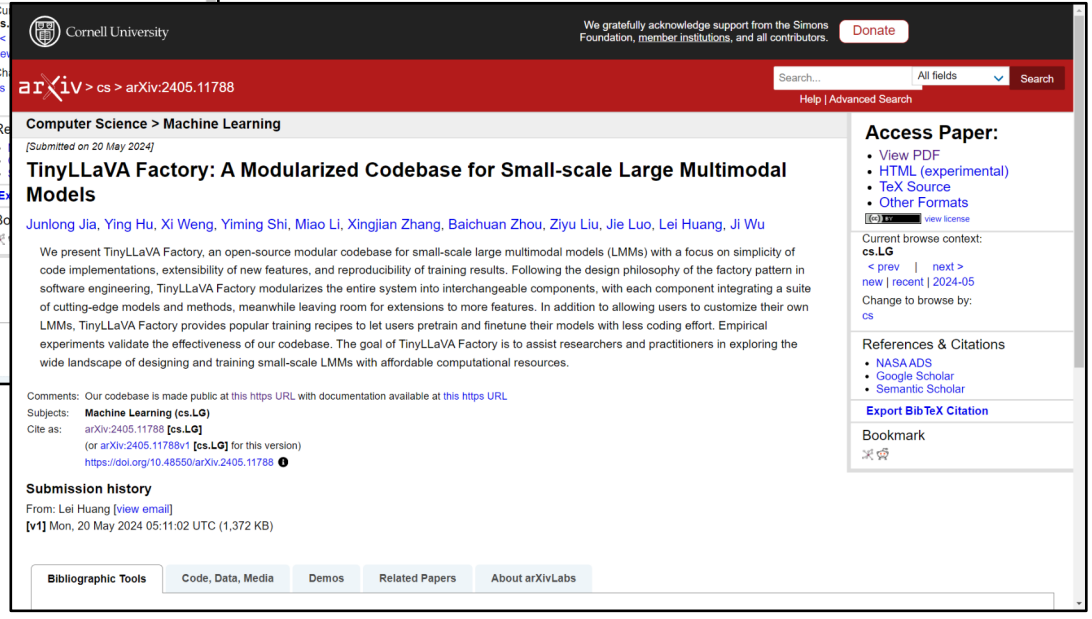
- Fine-Tune Your Own Tiny-Llama on Custom Dataset (YouTube)
- TinyLlama Colorist - fine-tuned with Color dataset (Medium)

PPO/DPO

- ?

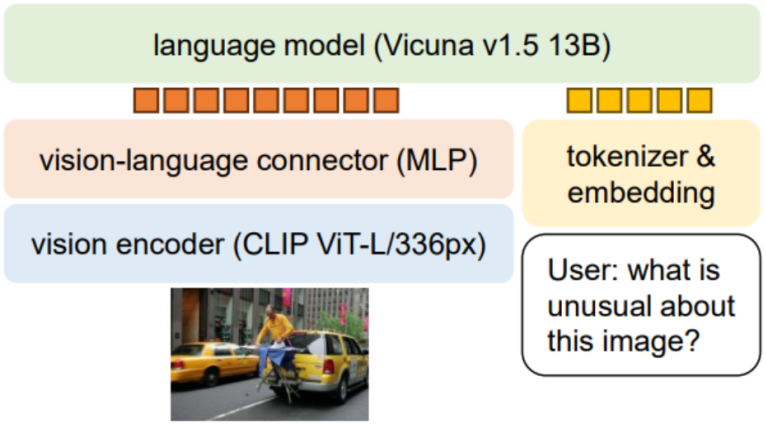
TinyLLaVA / TinyLLaVA Factory



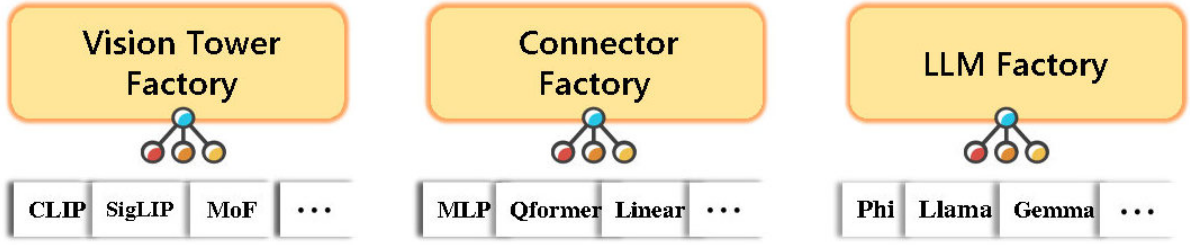


LLaVA / TinyLLaVA Factory

LLaVA



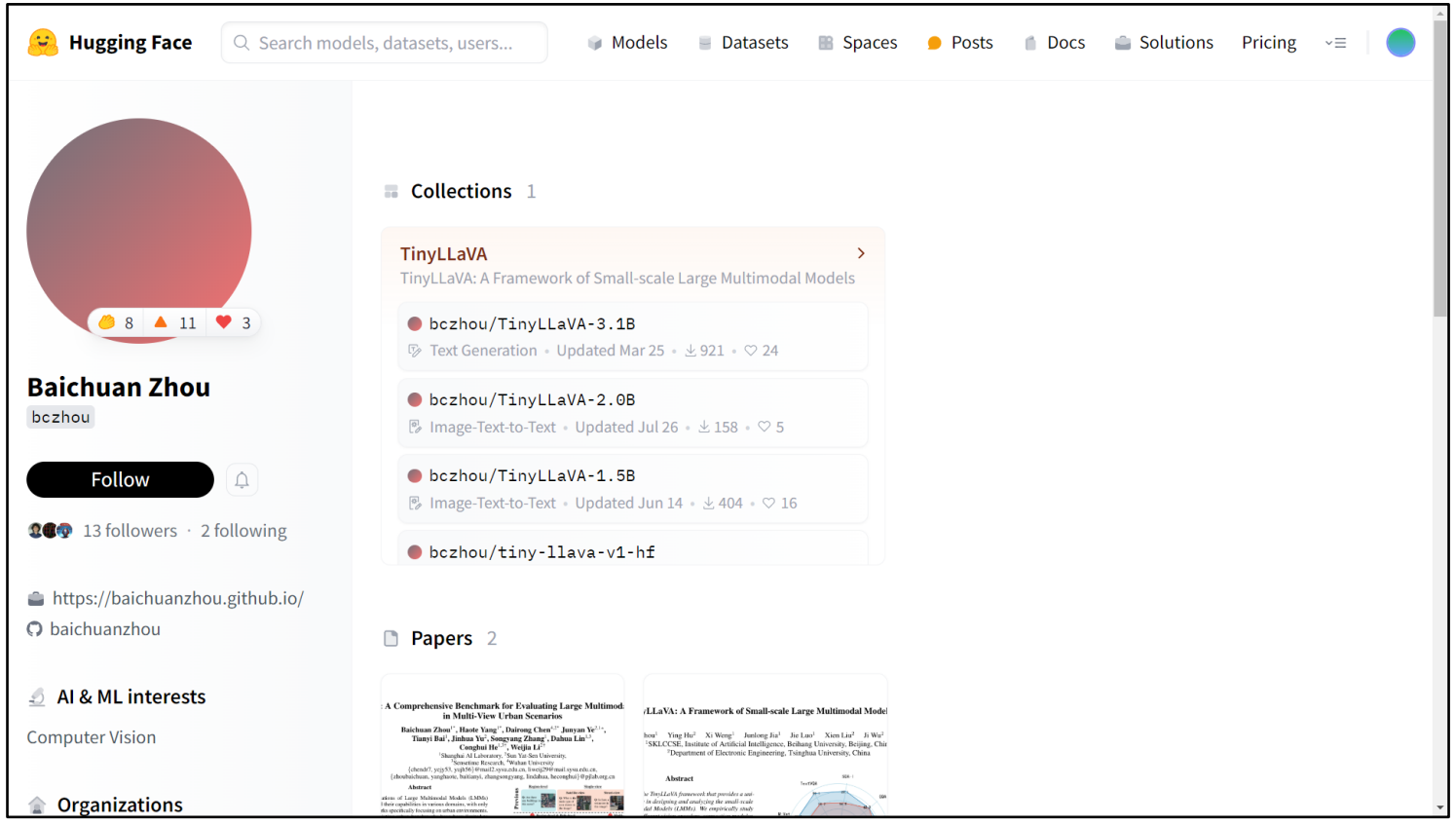
TinyLLaVA Factory



Module	Name	Module	Name
Small-scale LLM	OpenELM-450M (Mehta et al., 2024) TinyLlama-1.1B (Zhang et al., 2024) StableLM-2-1.6B (Jonathan Tow, 2023) Qwen-1.5-1.8B (Jinze Bai, 2024) Gemma-2B (Team et al., 2024) Phi-2-2.7B (Mojan Javaheripi, 2023)	Vision Tower	CLIP (Radford et al., 2021) SigLIP (Zhai et al., 2023) DinoV2 (Oquab et al., 2023) MoF (Tong et al., 2024)
Connector	Identity Linear MLP Q-former (Li et al., 2023) Resampler (Awadalla et al., 2023)	Training Recipe	Frozen/Full/Partially Tune LoRA/QLoRA (Hu et al., 2021; Dettmers et al., 2024)

https://github.com/TinyLLaVA/TinyLLaVA_Factory

TinyLLaVA Models



TinyLLaVA Factory : Finetuning & Customize

README Apache-2.0 license

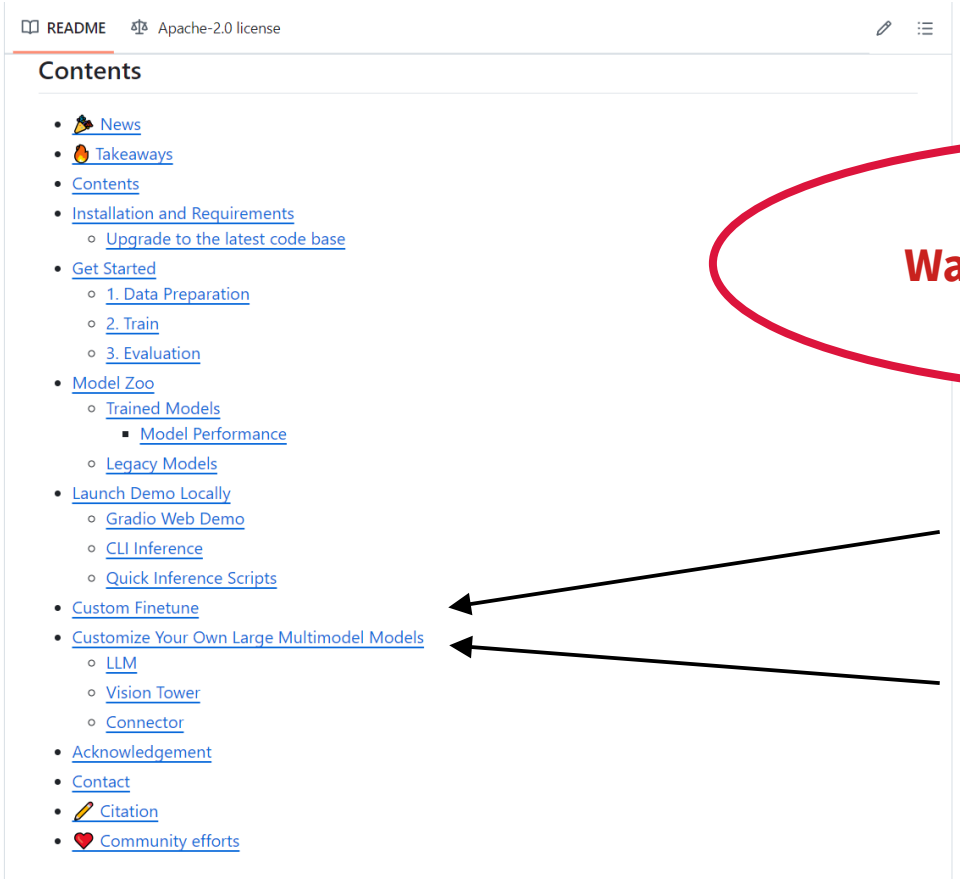
Contents

- News
- Takeaways
- Contents
- Installation and Requirements
  - Upgrade to the latest code base
- Get Started
  - 1. Data Preparation
  - 2. Train
  - 3. Evaluation
- Model Zoo
  - Trained Models
    - Model Performance
  - Legacy Models
- Launch Demo Locally
  - Gradio Web Demo
  - CLI Inference
  - Quick Inference Scripts
- Custom Finetune
- Customize Your Own Large Multimodel Models
  - LLM
  - Vision Tower
  - Connector
- Acknowledgement
- Contact
- Citation
- Community efforts

Supervised Finetuning
No PPO/DPO

Customize Components for New VLM

TinyLLaVA Factory : Finetuning & Customize


README Apache-2.0 license

Contents

- News
- Takeaways
- Contents
- Installation and Requirements
  - Upgrade to the latest code base
- Get Started
  - 1. Data Preparation
  - 2. Train
  - 3. Evaluation
- Model Zoo
  - Trained Models
    - Model Performance
  - Legacy Models
- Launch Demo Locally
  - Gradio Web Demo
  - CLI Inference
  - Quick Inference Scripts
- Custom Finetune
- Customize Your Own Large Multimodel Models
  - LLM
  - Vision Tower
  - Connector
- Acknowledgement
- Contact
- Citation
- Community efforts

Warning: Linux not Windows

Supervised Finetuning
No PPO/DPO

Customize Components for New VLM

TinyLLaMA, TinyLLaVA, TinyLLaVA Factory

余方國 博士
fangkuoyu@gmail.com