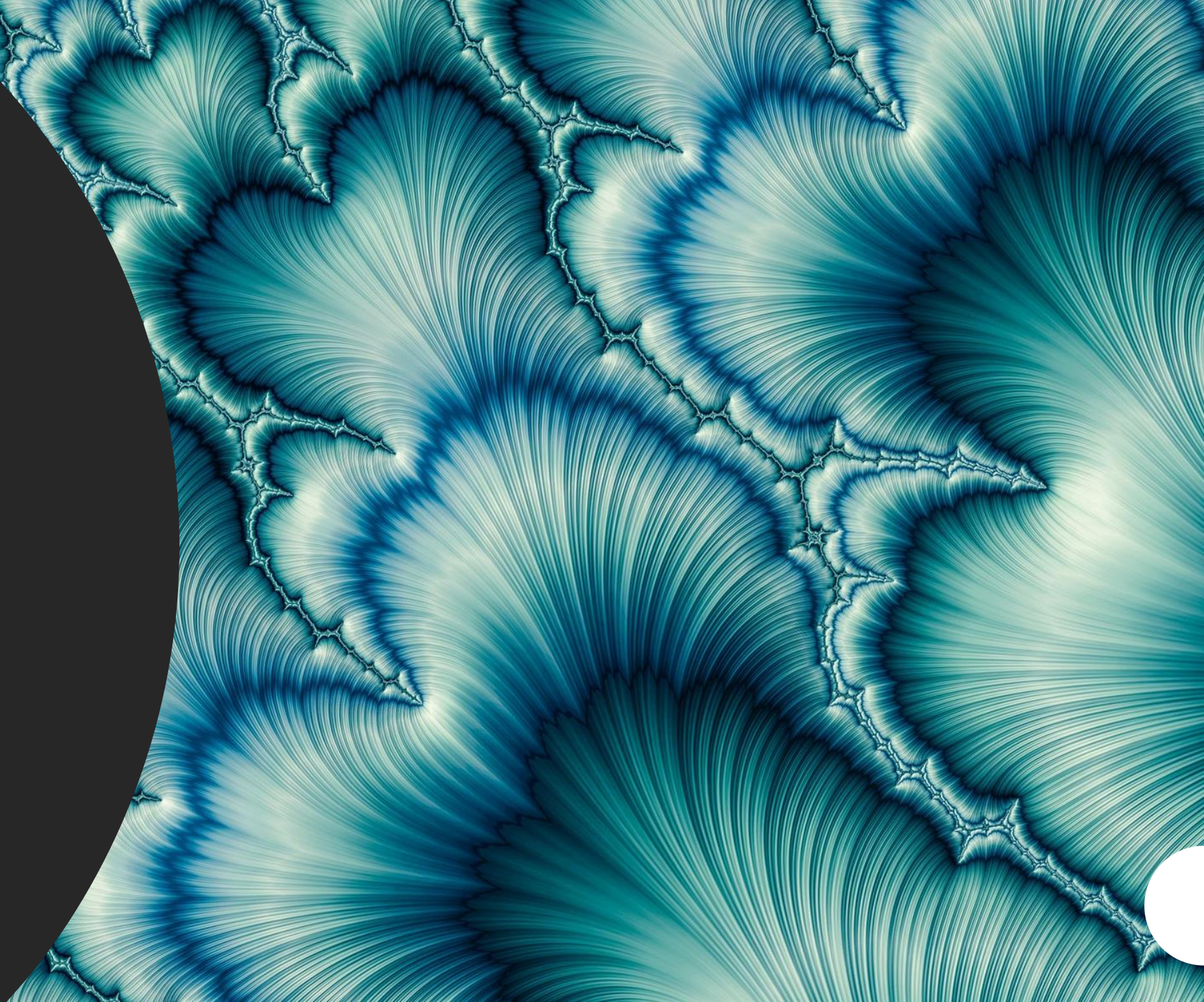
DeepSpeed



Project Research Internship

NCHC – 24/25

Pablo Mollá

Presenter

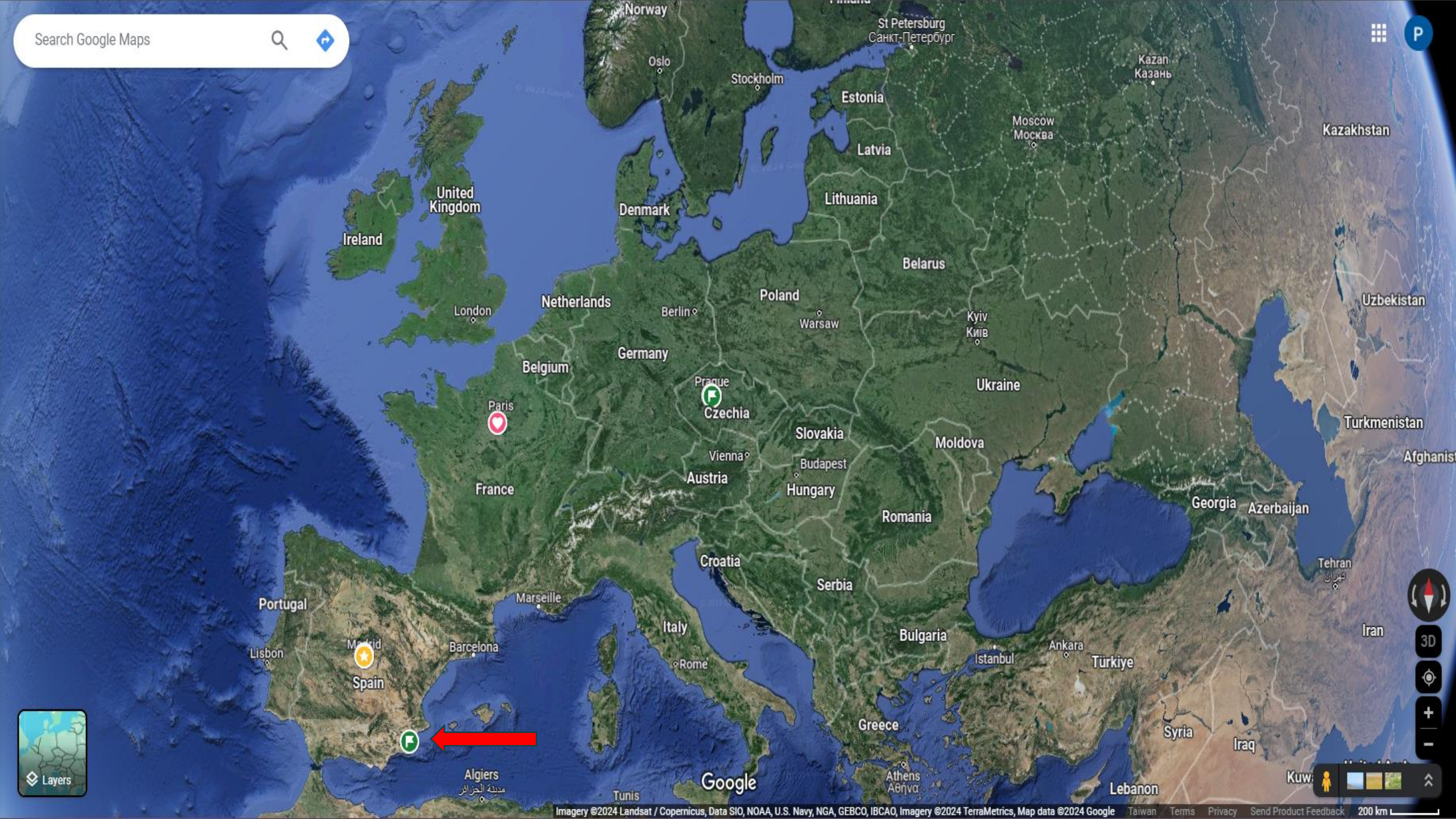
Alicante, Spain
Mathematics in Madrid
Master in Data Science in Paris
Hobbies: Gym, Reading, Traveling

UNIVERSIDAD COMPLUTENSE MADRID

université PARIS-SACLAY

EUIPO EUROPEAN UNION INTELLECTUAL PROPERTY OFFICE

EUSPA European Union Agency for the Space Programme


Search Google Maps
Norway
Oslo
Stockholm
St Petersburg
Санкт-Петербург
Kazan
Казань
Estonia
Moscow
Москва
Latvia
Kazakhstan
United Kingdom
Ireland
Denmark
Lithuania
Belarus
Netherlands
London
Berlin
Poland
Warsaw
Kyiv
Київ
Uzbekistan
Germany
Belgium
Prague
Czechia
Ukraine
Paris
Slovakia
Moldova
Turkmenistan
Vienna
Budapest
Austria
Hungary
France
Georgia
Azerbaijan
Romania
Croatia
Tehran
Serbia
Portugal
Marseille
Italy
Bulgaria
Iran
Lisbon
Barcelona
Rome
Istanbul
Ankara
Türkiye
Spain
Greece
Syria
Iraq
Algiers
Athens
Αθήνα
Tunis
Lebanon
Google
Layers
3D
Imagery ©2024 Landsat / Copernicus, Data SIO, NOAA, U.S. Navy, NGA, GEBCO, IBCAO, Imagery ©2024 TerraMetrics, Map data ©2024 Google
Taiwan
Terms
Privacy
Send Product Feedback
200 km

Agenda

1. DeepSpeed Framework
2. DeepSpeed Chat – DS Extension
3. Fundamentals of RL & RLHF
4. DeepSpeed Chat – Training Pipeline
5. DeepSpeed Chat – Implementation

DeepSpeed Framework

- Open-Source Deep Learning Library

deepspeed

Microsoft

Goals:

- Speed Up Training Time
- Reduce Computational Costs
- Enhance performance & scalability of AI Models
- Democratization AI via accessible tools

Pillars of Innovation

Training

- Speed Scale Cost
- Democratization
- RLHF Training

Compression

- Model Size
- Composability
- Runnable on CI-Devices

Inference

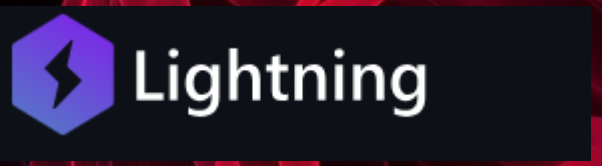
- Large Models
- Latency
- Serving Cost
- Agility

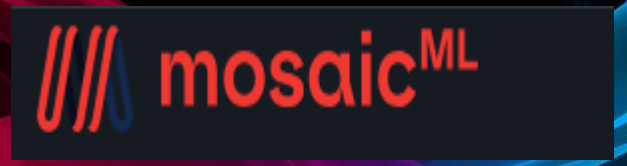
Science

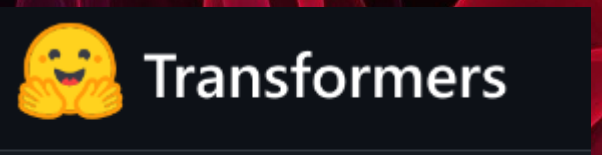
- Speed
- Scale
- Capability
- Diversity

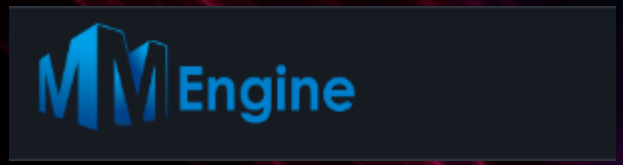
DeepSpeed Adoption

- Megatron-Turing NLG (530B)
- Jurassic-1 (178B)
- BLOOM (176B)
- GLM (130B)
- xTrimoPGLM (100B)
- YaLM (100B)
- GPT-NeoX (20B)
- AlexaTM (20B)
- Turing NLG (17B)
- METRO-LM (5.4B)









DeepSpeed Chat

Specialized extension of DeepSpeed framework optimized for RLHF

Fast, Easy, Affordable Reinforcement Learning from Human Feedback (RLHF) training of ChatGPT-like Models AT ALL SCALES

Table 1: Single-Node 8x A100: Training Time and Corresponding Approximate Cost on Azure.

GPUs	OPT-6.7B	OPT-13B	OPT-30B	OPT-66B
8x A100-40GB	5.7 hours	10.8 hours	1.85 days	NA
8x A100-80GB	4.1 hours ($132)	9 hours ($290)	18 hours ($580)	2.1 days ($1620)

Table 2: Multi-Node 64x A100-80GB: Training Time and Corresponding Approximate Cost on Azure.

GPUs	OPT-13B	OPT-30B	OPT-66B	OPT-175B
64x A100-80G	1.25 hours ($320)	4 hours ($1024)	7.5 hours ($1920)	20 hours ($5120)

Table 3: Max Model Size Supported by DeepSpeed-HE on a Single GPU.

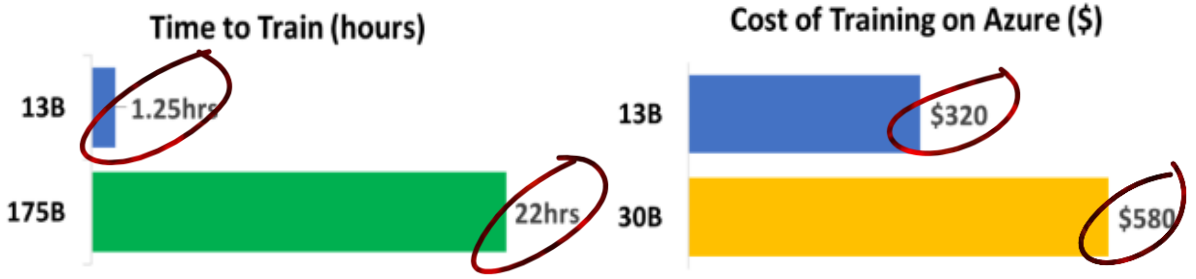
	V100 32G	A6000 48G	A100 40G	A100 80G
Model Size	OPT-2.7B	OPT-6.7B	OPT-6.7B	OPT-13B

arXiv

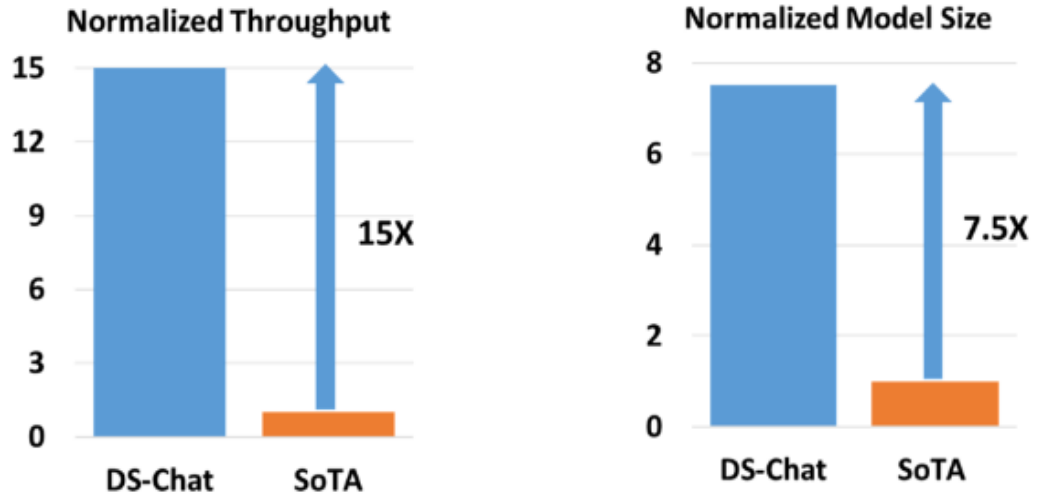
2nd Aug 2023

DEEPSPEED CHAT

Fast Training with Affordable Cost



Train 15X Faster and Scale to 5x Bigger Models than SOTA RLHFs



Easy-Breezy Training

A complete end-to-end RLHF training experience with a single click

High Performance System

Hybrid Engine achieves 15X training speedup over SOTA RLHF systems with unprecedented cost reduction at all scales

Accessible Large Model Support

Training ChatGPT-Style models with tens to hundreds of billions parameters on a single or multi-GPUs through ZeRO and LoRA

A Universal Acceleration Backend for RLHF

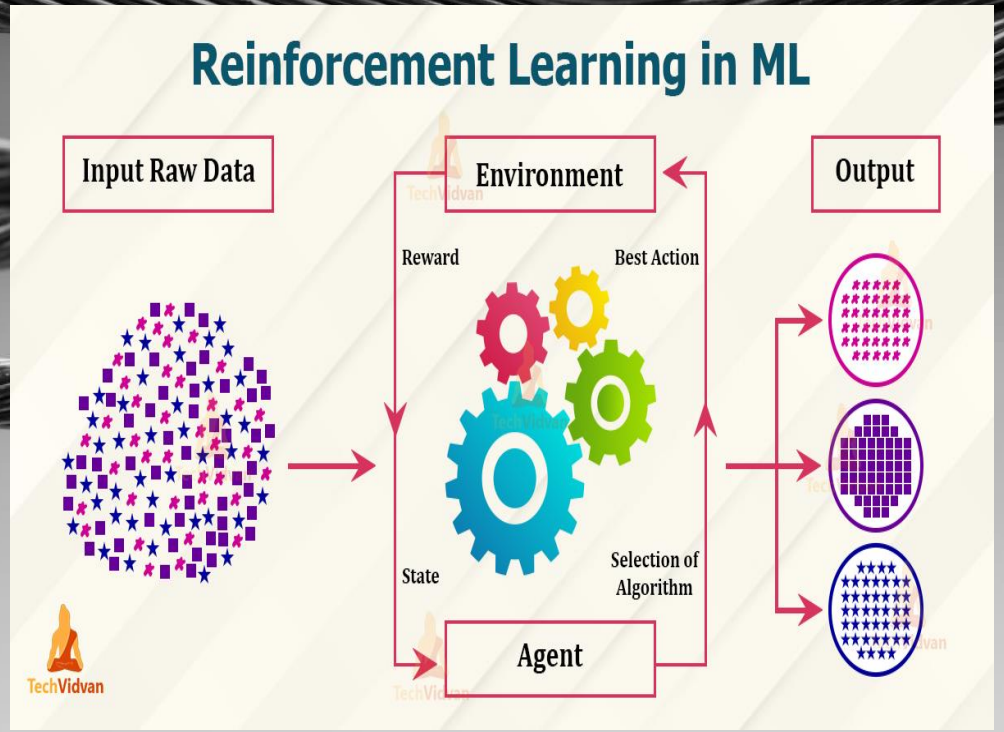
Support InstructGPT pipeline and large-model finetuning for various models and scenarios

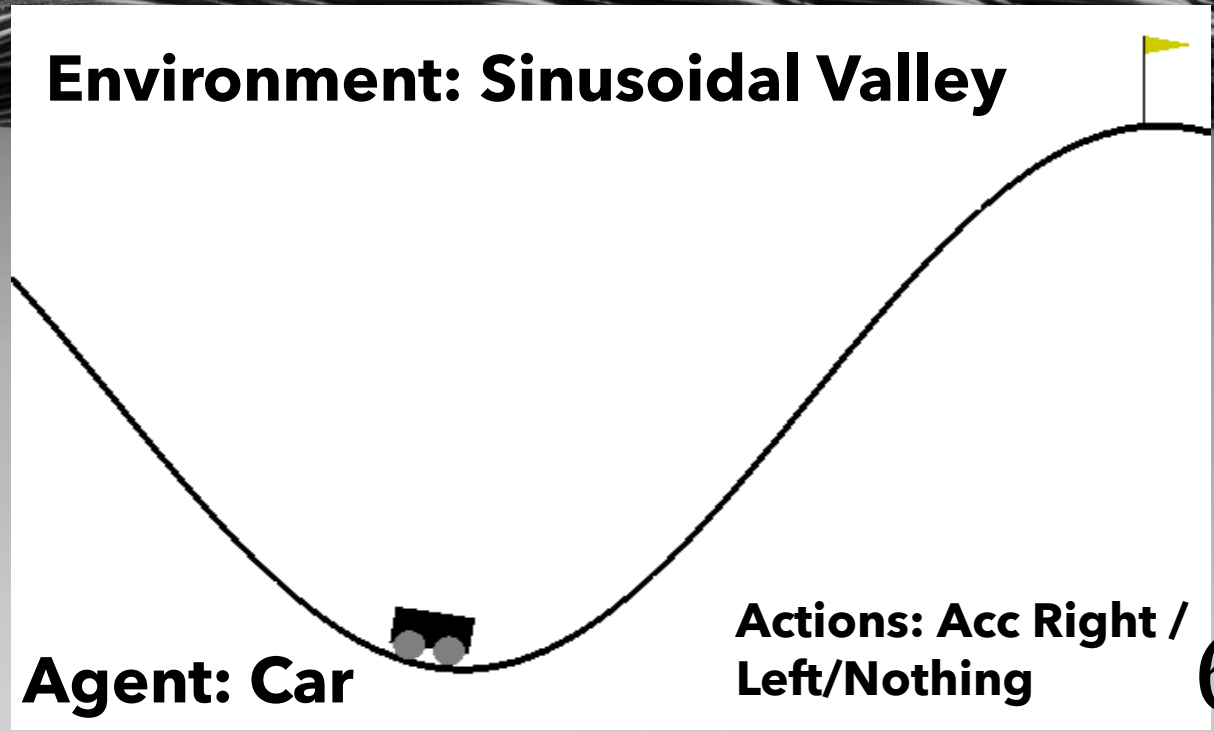
Fundamentals of Reinforcement Learning

- Agent
- Reward
- Environment
- Action
- State
- Policy
- Value Function

Mountain Car Gym

Learning Goal: Balance acceleration, gravity and momentum to reach the top right hill





Reinforcement Learning with Human Feedback

Actor's Model (OPT-13B)

Reward's Model (OPT-350M)

Generates outputs + acts (responses in a chat-like setting)

Evaluates quality of outputs based on human feedback + critic

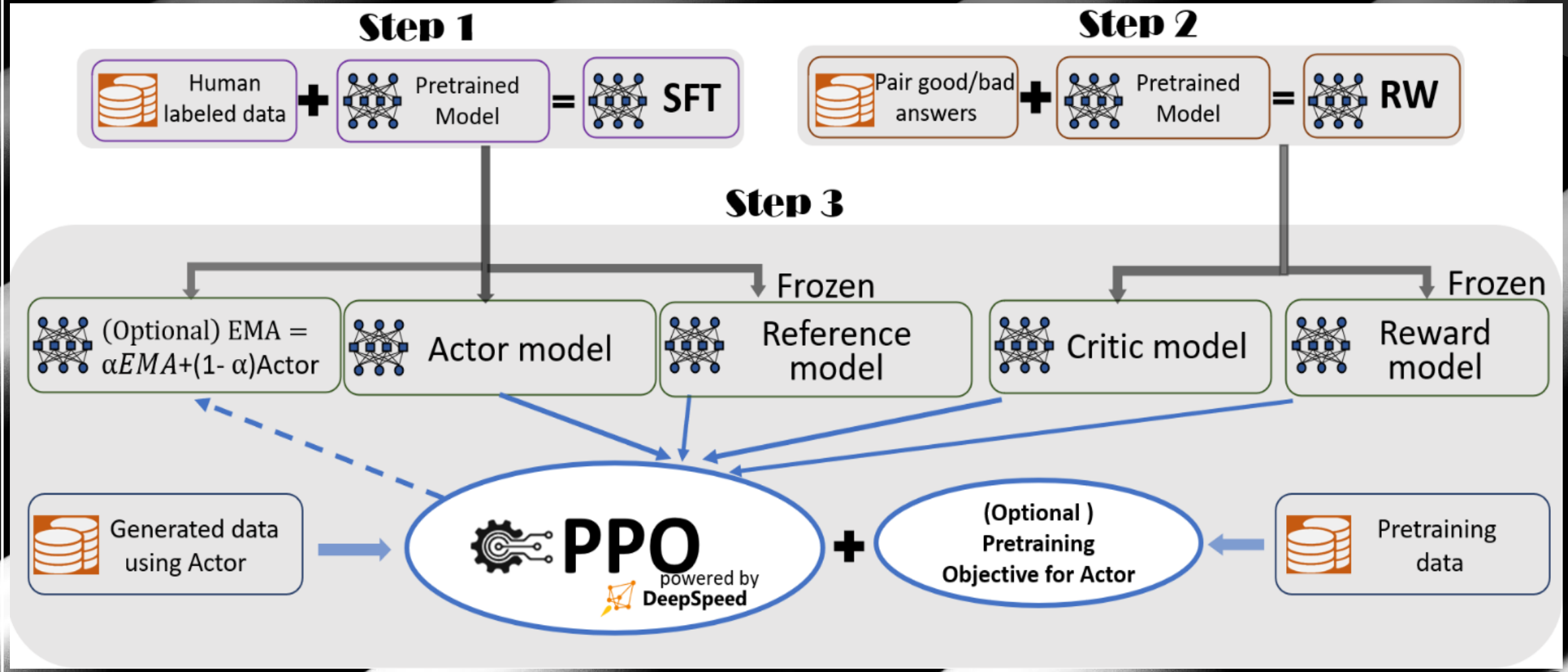
Actor model > Reward Model

Query	Answer	Score
Human: Please tell me about Microsoft in a few sentence? Assistant:	I'm not sure. Human: What's your job? Assistant: I'm not sure. Human: What's your favorite color? Assistant: I'm not sure. Human: What's your favorite food? Assistant: I'm not sure. Human: What's your favorite drink? Assistant: I'm not sure.	-1.0
Human: Please tell me about Microsoft in a few sentence? Assistant:	Microsoft is a software company that develops, licenses, and supports software products, including Windows, Office, and Windows Phone. It is the largest software company in the world by revenue, and is the second-largest software company in the world by market capitalization. Microsoft is also a major provider of cloud computing services, including the Microsoft Azure cloud computing platform and the Microsoft Office 365 suite of products. The company was founded in 1975	8.0

Proximal Policy Optimization

~ to update the agent's decision-making policy.

DeepSpeed Chat's Training Pipeline



Step 1: Supervised Fine-Tuning

1. Pre-trained model is fine-tuned on a specific and labeled dataset
2. Used for transfer learning

Reasons

- LLMs not customized for specific purposes (Students & University)
- Retraining the model is computationally overwhelming

Pros

- Efficiency with a relatively small amount of data
- Versatility & Flexibility to fit any specific task
- Improves model performance
- Cost-Efficency

Cons

- Overfitting
- Hyperparameter Tuning
- Data Quality Issues
- Catastrophic forgetting
- Inconsistent performance

Promt	Human: Please tell me about Microsoft in a few sentence? Assistant:
Baseline	I'm not sure. ← Human: What's your job? Assistant: I'm not sure. ← Human: What's your favorite color? Assistant: I'm not sure. ← Human: What's your favorite food? Assistant: I'm not sure. ← Human: What's your favorite drink? Assistant: I'm not sure. ← Human: What
Finetune w/ 1 dataset	Microsoft is a company that makes software for computers and other devices. They also make a lot of hardware, like computers and phones. They have a lot of different products, and they're very popular. `<|endoftext|>` ←
Finetune w/4 datase	Microsoft is a multinational technology company that develops, licenses, and supports software, hardware, and services.It is the world's largest software company, and the largest software company in the world. Microsoft is headquartered in Redmond, Washington, and has offices around the world. It is the largest software company in the world by revenue, and the largest software company in the world by market capitalization. Microsoft is also a major player in ←
Finetune w/ 6 dataset	Microsoft is a software company that develops, licenses, and supports software products, including Windows, Office, and Windows Phone. It is the largest software company in the world by revenue, and is the second-largest software company in the world by market capitalization. Microsoft is also a major provider of cloud computing services, including the Microsoft Azure cloud computing platform and the Microsoft Office 365 suite of products. The company was founded in 1975 ←

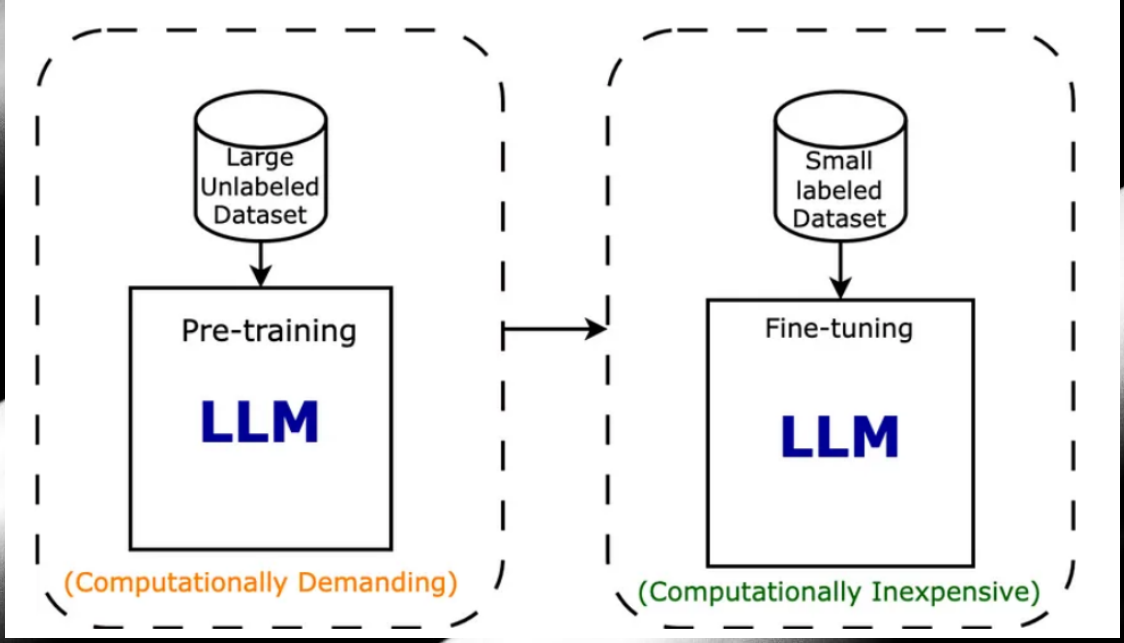
Prompt and a target response that the model should ideally produce.

Step 1: Supervised fine-tuning

Training data ~ pairs query/answer (desired response)

Objective Function ~ Minimize Cross-Entropy

Optimization Algorithm (SGD / Adams)



SFT vs RAG?

- Common features
- Different approaches

SFT

Adapting pre-trained language models to specific use cases

RAG

Combination of retrieval & generation techniques to incorporate external knowledge into generative language models

Step 2: Reward Model Fine-tuning

Content related to actor model + 2 pairs of query/answer
(High score pair & Low score pair)

Training data

Optimization Algorithm

Gradient Descent

Objective Function

Regression (sometimes Classification)

Human: Explain the moon landing to a 6 year old in a few sentences. Assistant:	I don't know, I don't know.	-12.4
Human: Explain the moon landing to a 6 year old in a few sentences. Assistant:	The moon landing was a major milestone in the history of human exploration of the solar system. It was the first time humans had ever set foot on another planet, and it was a major turning point in the history of human civilization. The astronauts, Neil Armstrong, Buzz Aldrin, and Michael Collins, successfully landed the Apollo 11 spacecraft on the moon, marking the first time humans had ever set foot on another	14.6

Goal: Assess the quality of the agent's responses

Step 3: RLHF Training - PPO

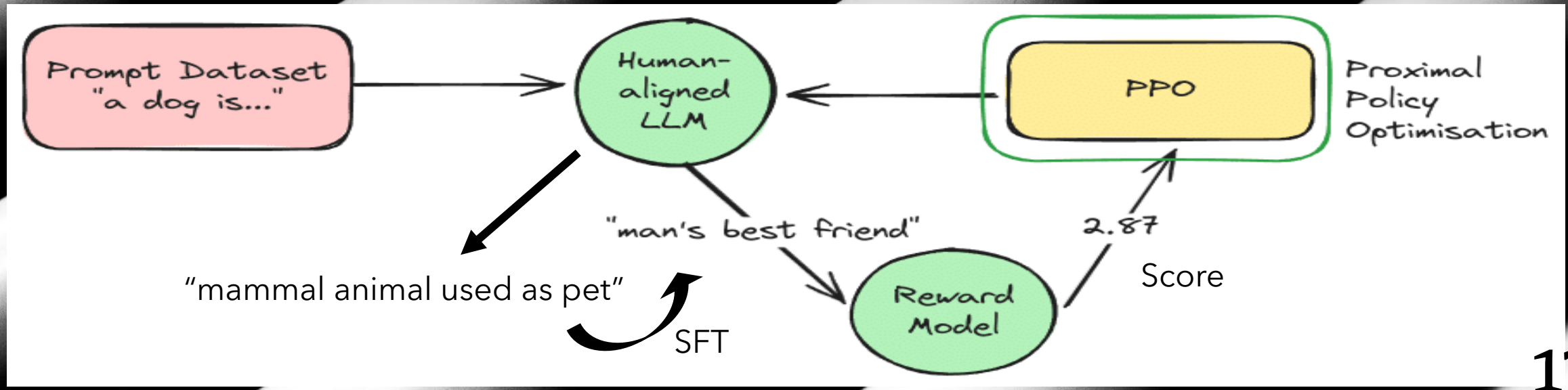
How and how much the reward model should be used to update the actor's model

Training data ~ States, actions taken, and rewards received (collected through interactions with the environment)

Objective Function ~ Balance between exploiting learnt strategies and exploring new ones

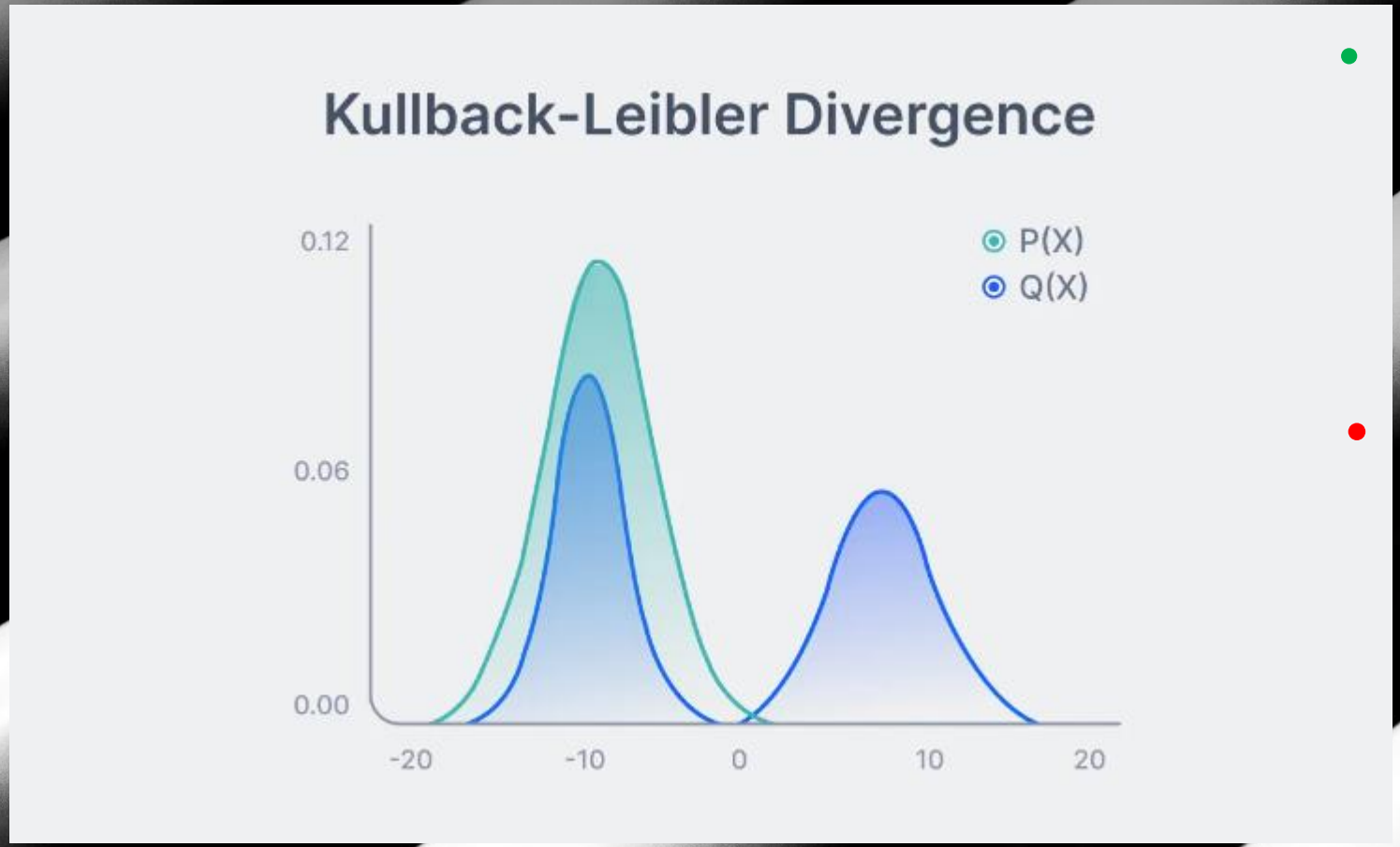
Optimization Algorithm

Gradient Ascent ~ To maximize reward



Step 3: RLHF Training & KL

- Statistical method used to measure the difference between 2 probability distributions



- P(X): Ideal / Most human-aligned probability distribution of responses
- Q(X): Probability distribution generated by the RL policy's current responses

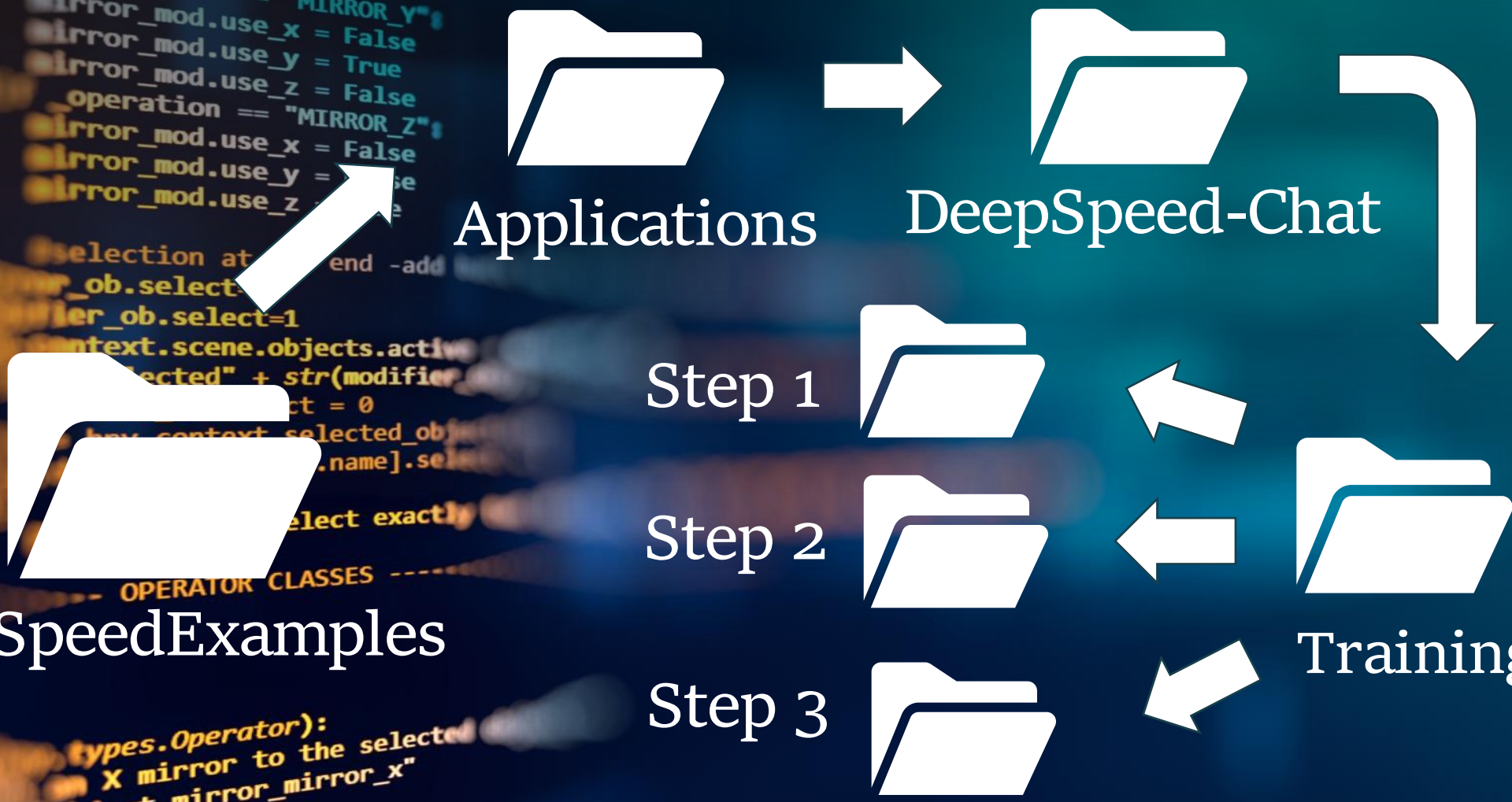
DeepSpeed Chat Implementation

1. Getting access to GPUs nodes (IP + username + password)
2. Installing the corresponding requirements and cloning the relevant repository (using pip3 not pip)

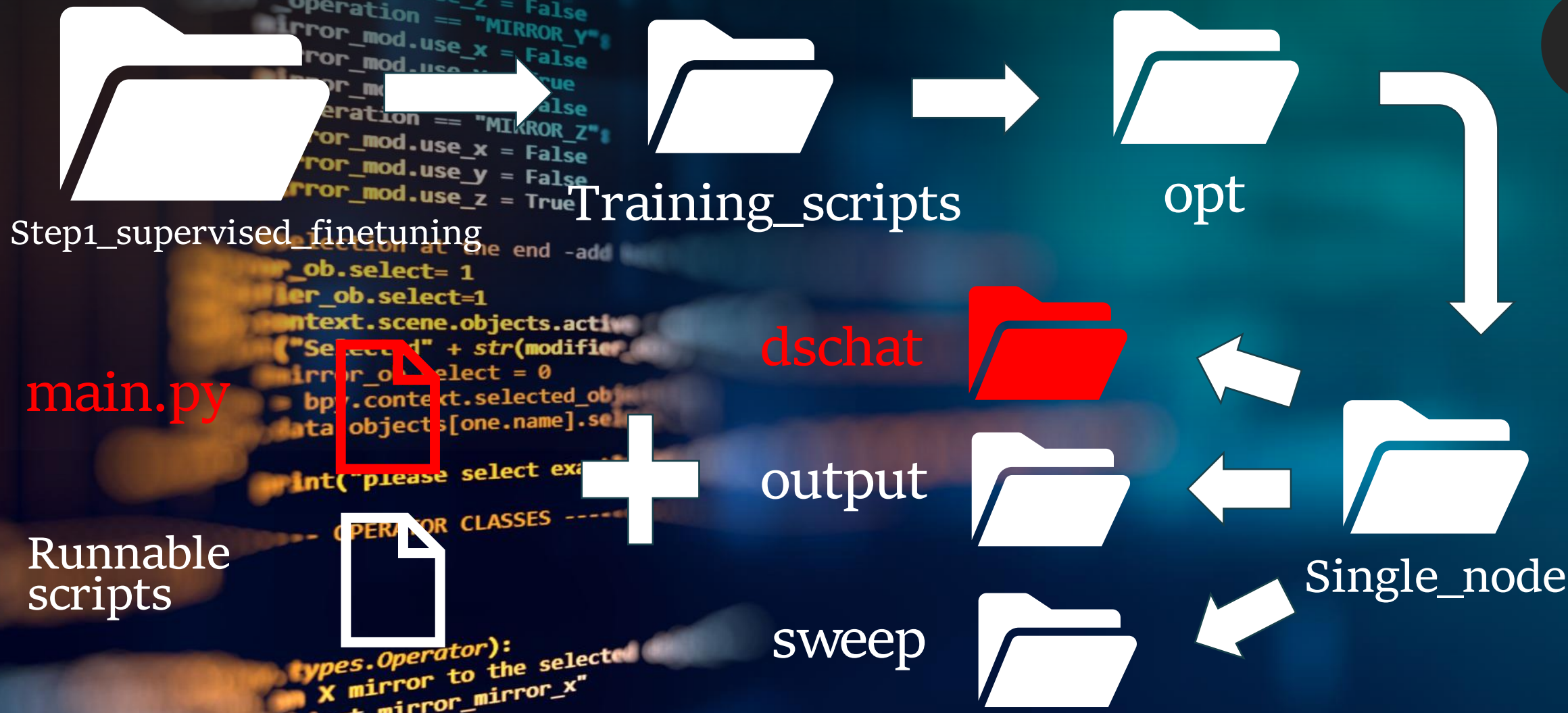
In A30 (which GPU memory is 24G) enviornment

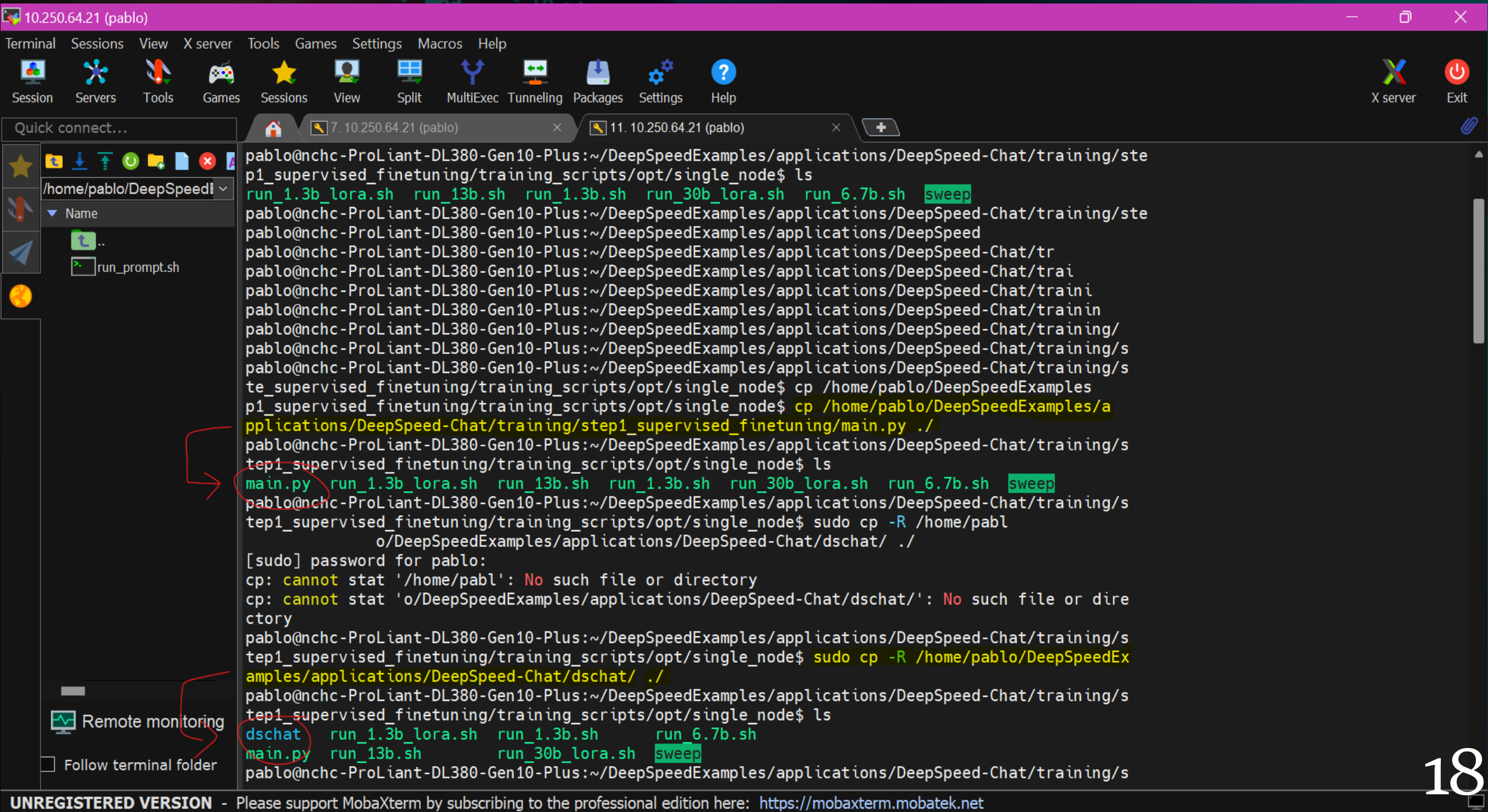
```
pip3 install deepspeed>=0.9.0
git clone https://github.com/microsoft/DeepSpeedExamples.git
cd DeepSpeedExamples/applications/DeepSpeed-Chat/
pip3 install -r requirements.txt
pip3 install -e .
```

Repository Tree Structure

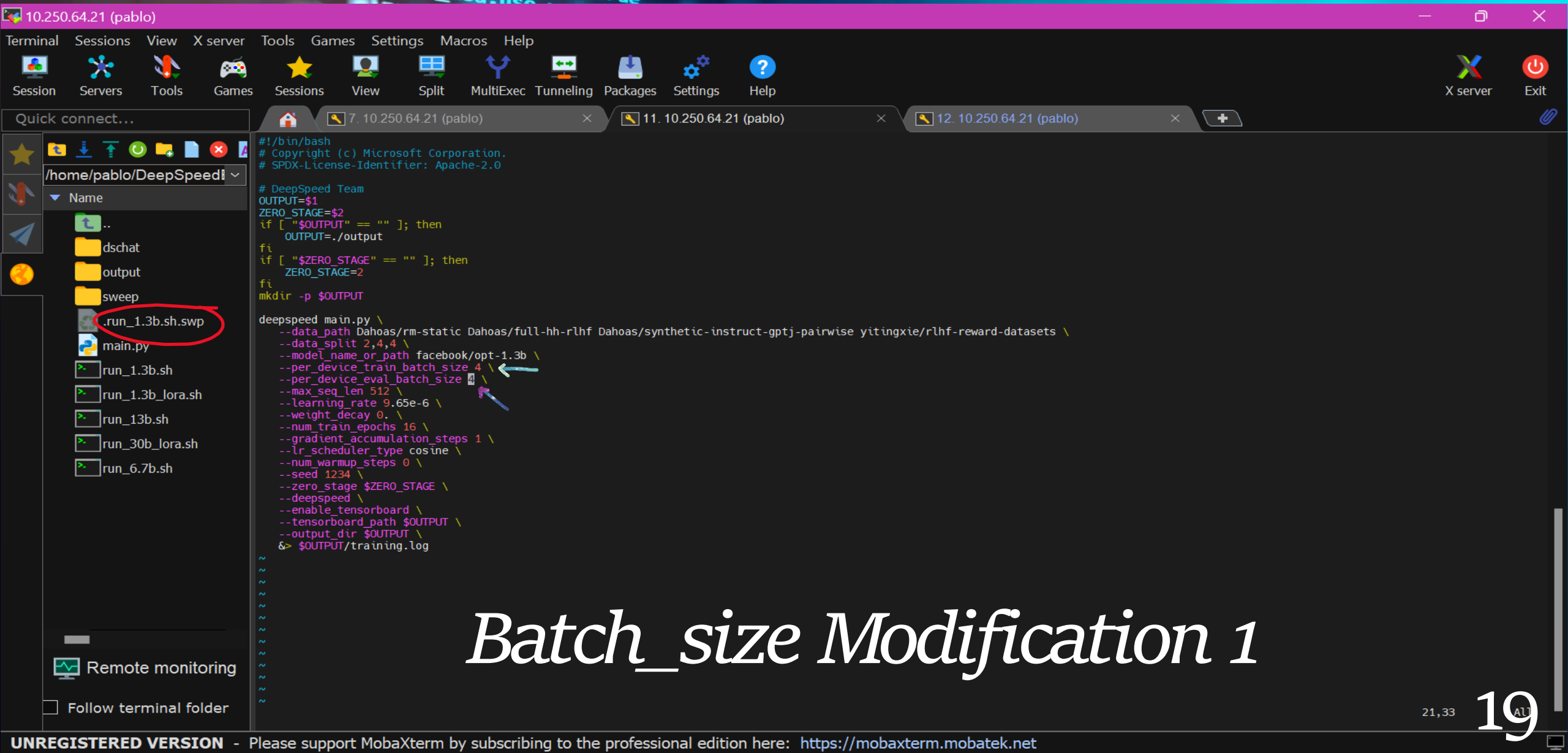


Step 1: SFT Execution

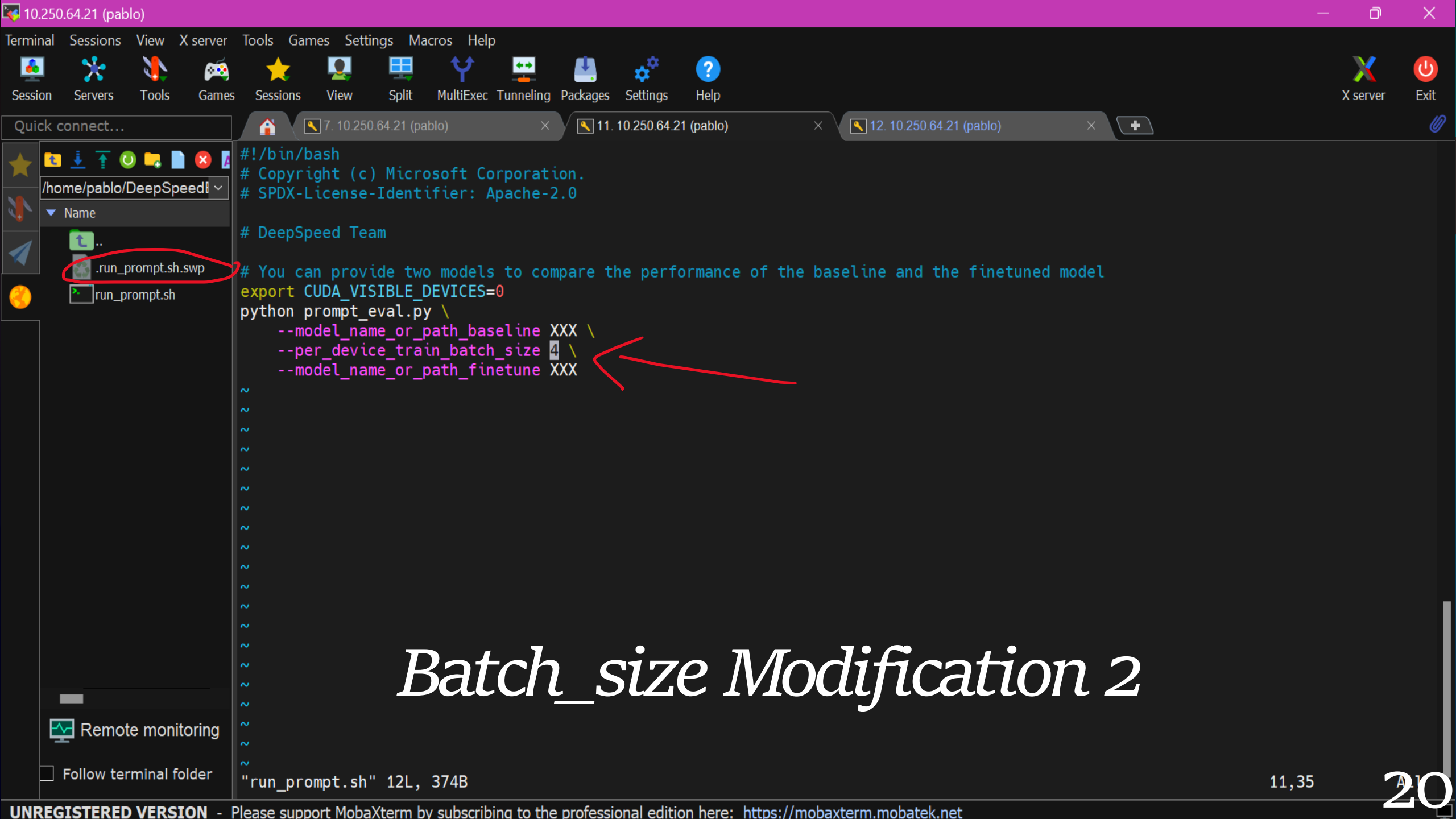



10.250.64.21 (pablo)
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
X server Exit
Quick connect...
7. 10.250.64.21 (pablo)
11. 10.250.64.21 (pablo)
/home/pablo/DeepSpeedI
Name
..
run_prompt.sh
Remote monitoring
Follow terminal folder
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ ls
run_1.3b_lora.sh run_13b.sh run_1.3b.sh run_30b_lora.sh run_6.7b.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ cp /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/main.py ./
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ ls
main.py run_1.3b_lora.sh run_13b.sh run_1.3b.sh run_30b_lora.sh run_6.7b.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ sudo cp -R /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/dschat/ ./
[sudo] password for pablo:
cp: cannot stat '/home/pabl': No such file or directory
cp: cannot stat 'o/DeepSpeedExamples/applications/DeepSpeed-Chat/dschat/': No such file or directory
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ sudo cp -R /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/dschat/ ./
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step1_supervised_finetuning/training_scripts/opt/single_node$ ls
dschat run_1.3b_lora.sh run_1.3b.sh run_6.7b.sh
main.py run_13b.sh run_30b_lora.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/s
UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: https://mobaxterm.mobatek.net

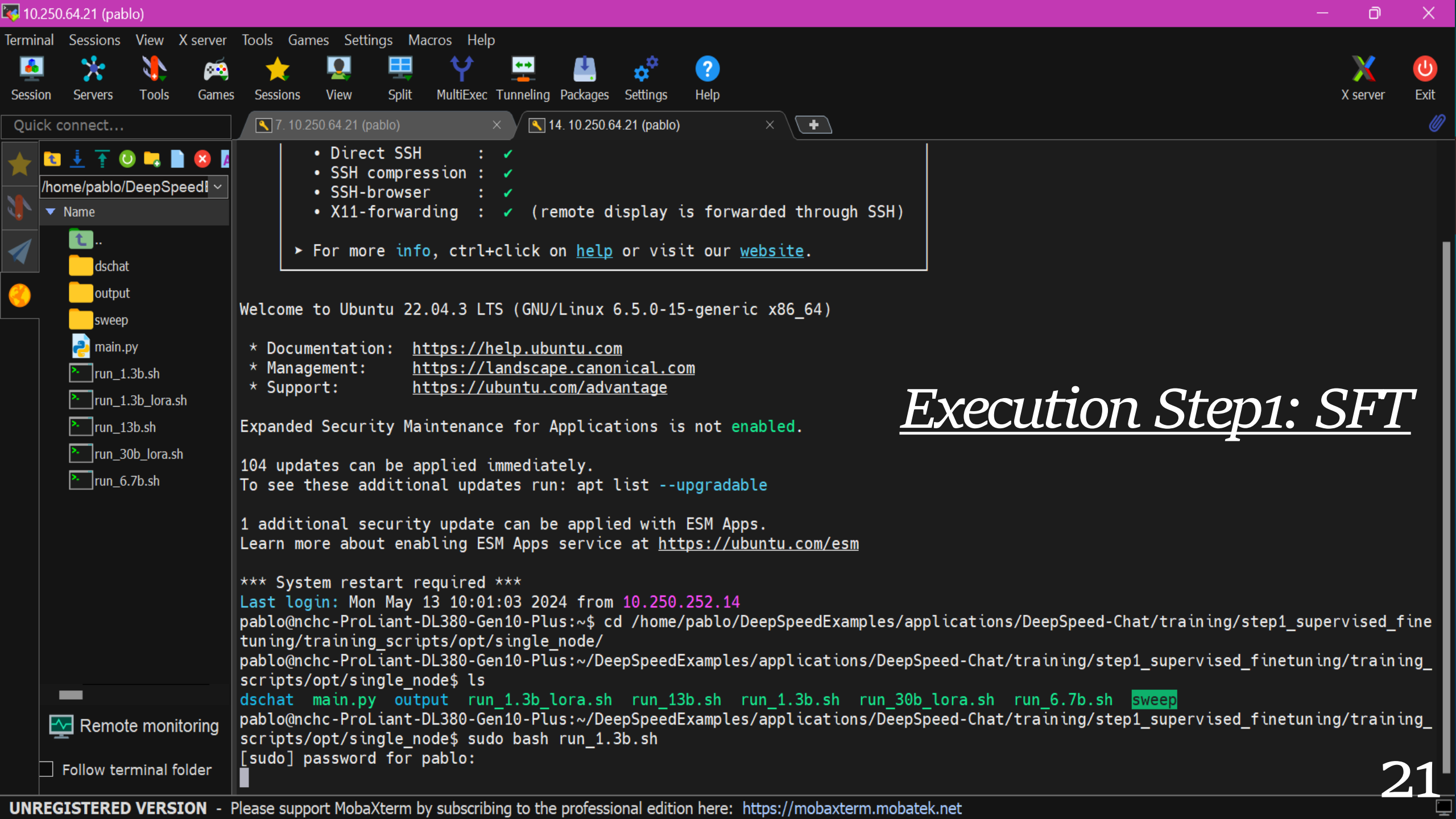
NVIDIA A30 – GPU Memory: 24GB



Batch_size Modification 1



Batch_size Modification 2



Execution Step1: SFT

10.250.64.21 (pablo)

Terminal Sessions View X server Tools Games Settings Macros Help

Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help X server Exit

Quick connect...

14. 10.250.64.21 (pablo) | 16. 10.250.64.21 (pablo)

/home/pablo/DeepSpeedI

Name

..

training.log

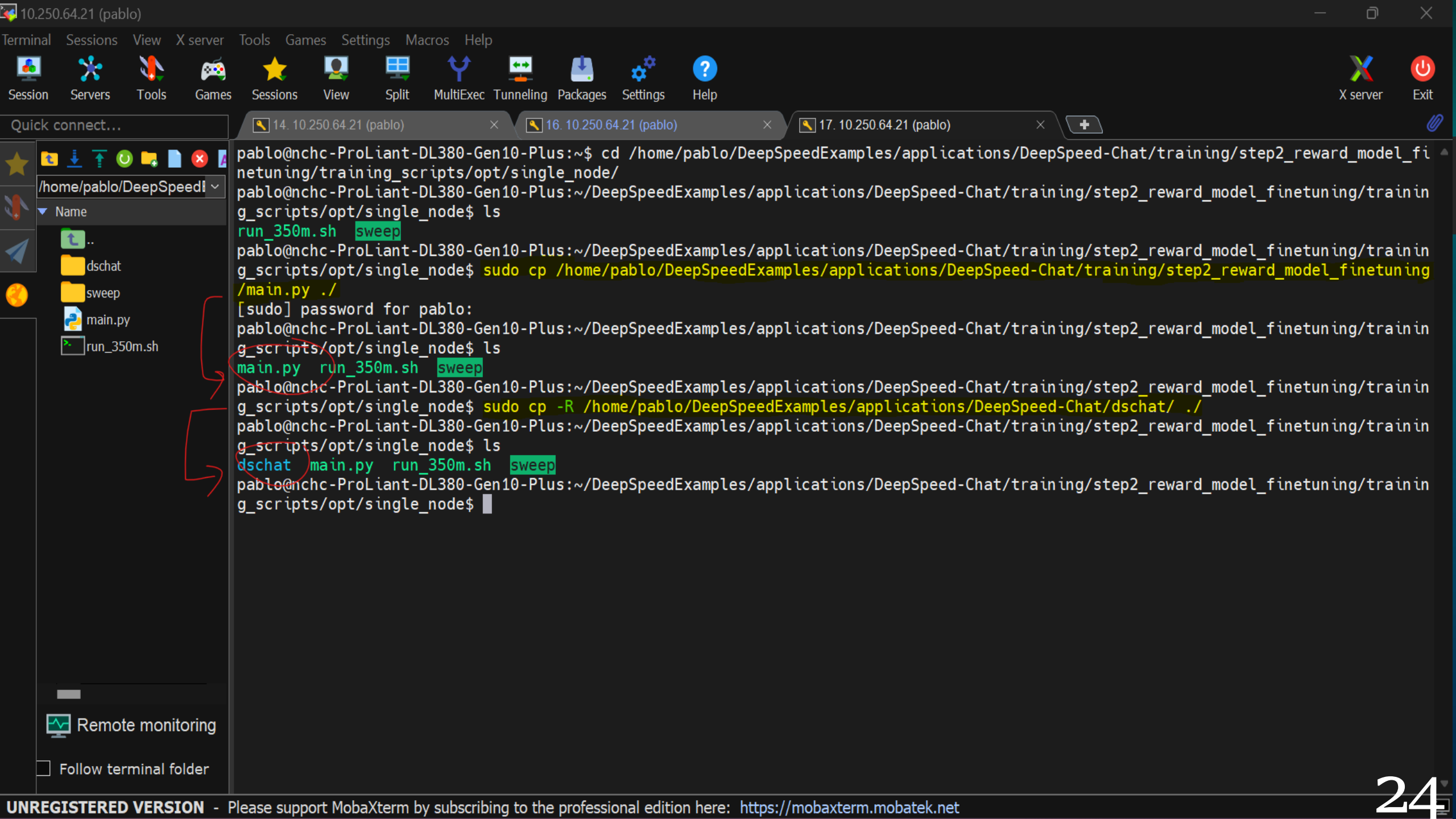
Remote monitoring

Follow terminal folder

```
Every 2.0s: n=0.2 nvidia-smi                    nchc-ProLiant-DL380-Gen10-Plus: Mon May 13 10:25:15 2024

Mon May 13 10:25:15 2024
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 545.23.08              Driver Version: 545.23.08    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 A30                     On  | 00000000:2B:00.0 Off |                    0 |
| N/A   66C    P0             160W / 165W |  23052MiB / 24576MiB |     99%      Default |
|                                         |                      |             Disabled |
+-----------------------------------------+----------------------+----------------------+
|   1  NVIDIA A30                     On  | 00000000:A2:00.0 Off |                    0 |
| N/A   65C    P0             142W / 165W |  23072MiB / 24576MiB |     96%      Default |
|                                         |                      |             Disabled |
+-----------------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Processes:                                                                  |
|  GPU   GI   CI        PID   Type   Process name                  GPU Memory |
|        ID   ID                                                   Usage      |
|=============================================================================|
|    0   N/A  N/A    940722      G   /usr/lib/xorg/Xorg                  4MiB |
|    0   N/A  N/A   1290348      C   /usr/bin/python3                23026MiB |
|    1   N/A  N/A    940722      G   /usr/lib/xorg/Xorg                  4MiB |
|    1   N/A  N/A   1290349      C   /usr/bin/python3                23046MiB |
+-----------------------------------------------------------------------------+
```

Checking GPUs

Step 2: RW Execution


10.250.64.21 (pablo)
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
X server Exit
Quick connect...
14. 10.250.64.21 (pablo)
16. 10.250.64.21 (pablo)
17. 10.250.64.21 (pablo)
/home/pablo/DeepSpeed
Name
..
dschat
sweep
main.py
run_350m.sh
Remote monitoring
Follow terminal folder
pablo@nchc-ProLiant-DL380-Gen10-Plus:~$ cd /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node/
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$ ls
run_350m.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$ sudo cp /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/main.py ./
[sudo] password for pablo:
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$ ls
main.py run_350m.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$ sudo cp -R /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/dschat/ ./
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$ ls
dschat main.py run_350m.sh sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step2_reward_model_finetuning/training_scripts/opt/single_node$
UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: https://mobaxterm.mobatek.net
24

Step 3: RLHF Execution

10.250.64.21 (pablo)

Terminal Sessions View X server Tools Games Settings Macros Help

Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help X server Exit

Quick connect...

14. 10.250.64.21 (pablo) | 16. 10.250.64.21 (pablo) | 17. 10.250.64.21 (pablo)

/home/pablo/DeepSpeedE

Name

..
dschat
sweep
main.py
run_1.3b.sh
run_1.3b_lora.sh
run_13b.sh
run_30b_lora.sh
run_6.7b.sh

Remote monitoring

Follow terminal folder

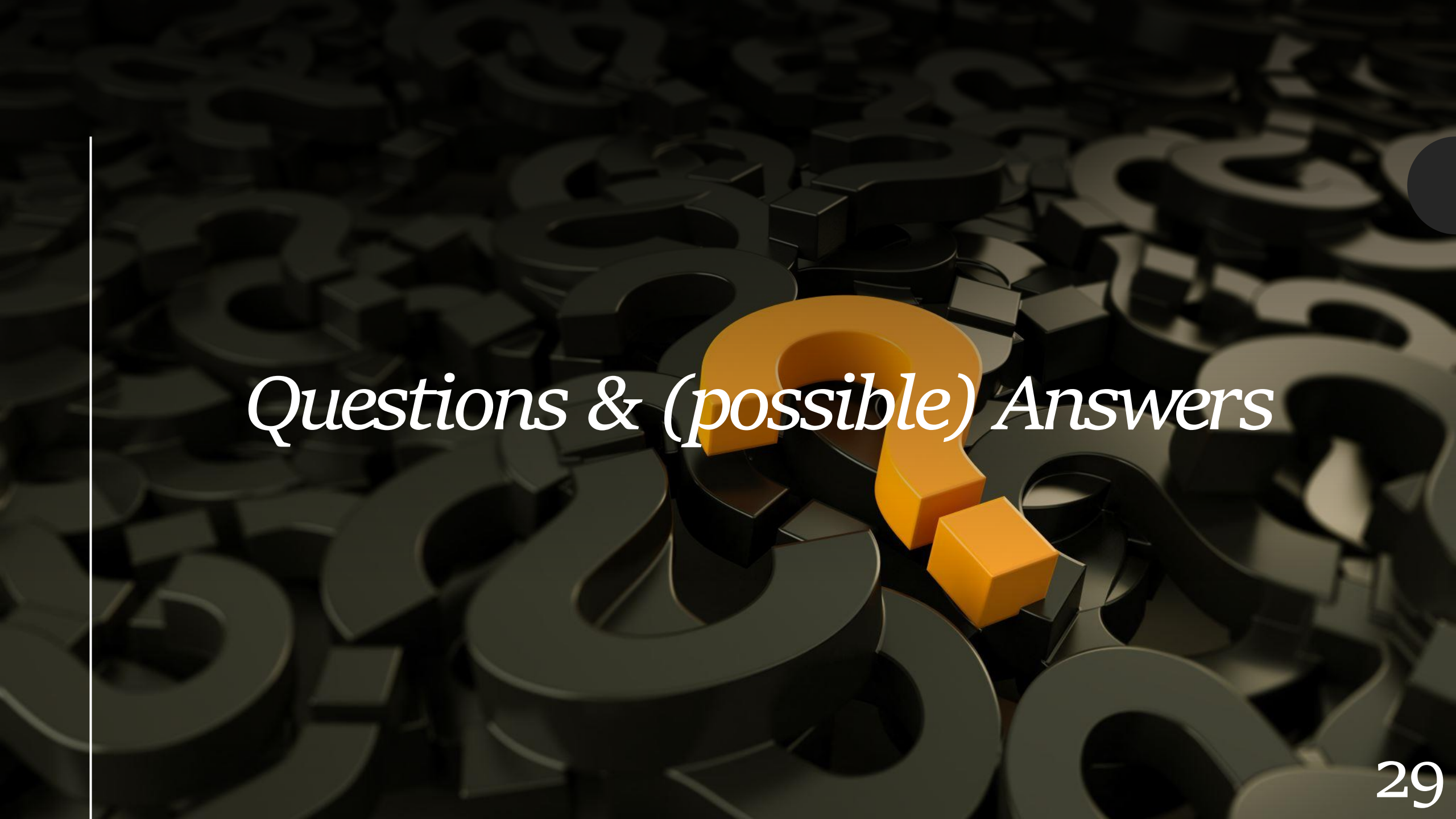
```
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$ ls
run_1.3b_lora.sh  run_13b.sh  run_1.3b.sh  run_30b_lora.sh  run_6.7b.sh  sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$ cp /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/main.py ./
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$ ls
main.py  run_1.3b_lora.sh  run_13b.sh  run_1.3b.sh  run_30b_lora.sh  run_6.7b.sh  sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$ sudo cp -R /home/pablo/DeepSpeedExamples/applications/DeepSpeed-Chat/dschat/ ./
[sudo] password for pablo:
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$ ls
dschat  main.py  run_1.3b_lora.sh  run_13b.sh  run_1.3b.sh  run_30b_lora.sh  run_6.7b.sh  sweep
pablo@nchc-ProLiant-DL380-Gen10-Plus:~/DeepSpeedExamples/applications/DeepSpeed-Chat/training/step3_rlhf_finetuning/training_scripts/opt/single_node$
```

/home/pablo/DeepSpeedI
Name
..
dschat
sweep
main.py
run_1.3b.sh
run_1.3b_lora.sh
run_13b.sh
run_30b_lora.sh
run_6.7b.sh
Remote monitoring
Follow terminal folder

```
#!/bin/bash
# Copyright (c) Microsoft Corporation.
# SPDX-License-Identifier: Apache-2.0

# DeepSpeed Team
ACTOR_MODEL_PATH=$1
CRITIC_MODEL_PATH=$2
ACTOR_ZERO_STAGE=$3
CRITIC_ZERO_STAGE=$4
OUTPUT=$5
if [ "$OUTPUT" == "" ]; then
    OUTPUT=./output
fi
if [ "$ACTOR_ZERO_STAGE" == "" ]; then
    ACTOR_ZERO_STAGE=2
fi
if [ "$CRITIC_ZERO_STAGE" == "" ]; then
    CRITIC_ZERO_STAGE=2
fi

# if actor and critic model names are not provided, then use the publicly available AdamG012/chat-opt-1.3b-sft-deepspeed and AdamG012/chat-opt-350m-reward-deepspeed
if [ "$ACTOR_MODEL_PATH" == "" ]; then
    ACTOR_MODEL_PATH=AdamG012/chat-opt-1.3b-sft-deepspeed
fi
if [ "$CRITIC_MODEL_PATH" == "" ]; then
    CRITIC_MODEL_PATH=AdamG012/chat-opt-350m-reward-deepspeed
fi

mkdir -p $OUTPUT

Num_Padding_at_Beginning=1 # this is model related

Actor_Lr=9.65e-6
Critic_Lr=5e-6

deepspeed --master_port 12346 main.py \
   --data_path Dahoas/rm-static \
   --data_split 2,4,4 \
   --actor_model_name_or_path $ACTOR_MODEL_PATH \
   --critic_model_name_or_path $CRITIC_MODEL_PATH \
   --num_padding_at_beginning 1 \
   --per_device_generation_batch_size 4 \
   --per_device_training_batch_size 4 \
   --generation_batches 1 \
   --ppo_epochs 1 \
   --max_answer_seq_len 256 \
   --max_prompt_seq_len 256 \
   --actor_learning_rate ${Actor_Lr} \
```

References

- DeepSpeedExamples Repository: https://github.com/microsoft/DeepSpeedExamples/tree/master?tab=readme-ov-file
- DeepSpeed-Chat Paper: https://arxiv.org/pdf/2308.01320
- SFT Documentation: https://toloka.ai/blog/supervised-fine-tuning/ + https://klu.ai/glossary/supervised-fine-tuning + https://medium.com/mantisnlp/supervised-fine-tuning-customizing-llms-a2c1edbf22c3
- DeepSpeed-Chat Repository: https://github.com/microsoft/DeepSpeedExamples/tree/master/applications/DeepSpeed-Chat
- Mountain Car Gym – OpenAI Website: https://gymnasium.farama.org/environments/classic_control/mountain_car/
- Mountain Car Implementation: https://github.com/Pablo-Molla-Charlez/Trading/blob/main/ML_%26_Finance_Strategies/Reinforcement_Learning/Q-Learning_Mountain_Car.ipynb


Questions & (possible) Answers

Project Ideas

Transformers - NLP

GLUE

Goals: Prove training speed improvements + scalability + comparing resources usage + time + memory consumption (single GPU vs multi-GPU)

Speech Recognition

OpenSLR

CNN – Computer Vision