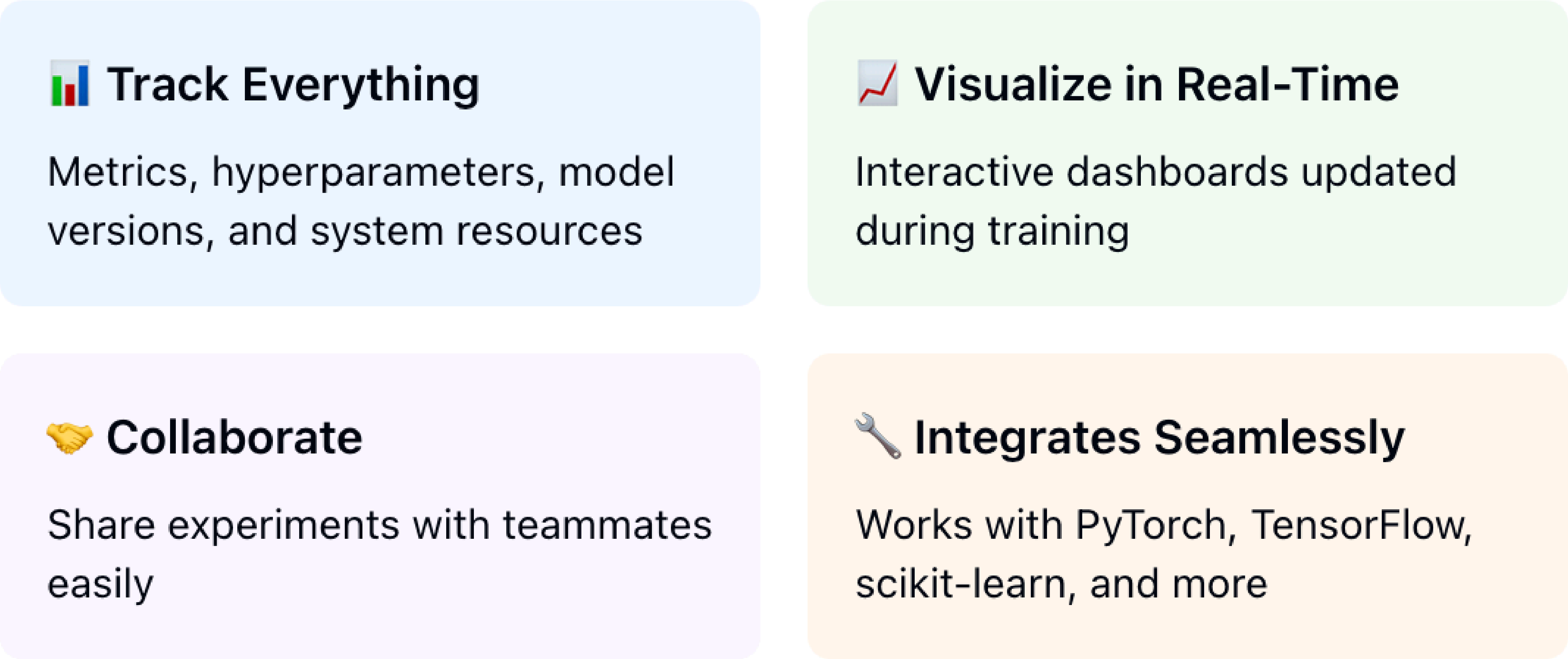
INTRODUCTION TO WANDB

Weights & Biases

By: N.S.Akshara

WHAT IS WANDB?

WandB (Weights and Biases) is a powerful ML Operations platform that helps data scientists and ML engineers track, visualize, and manage their experiments seamlessly.



📊 Track Everything

Metrics, hyperparameters, model versions, and system resources

📈 Visualize in Real-Time

Interactive dashboards updated during training

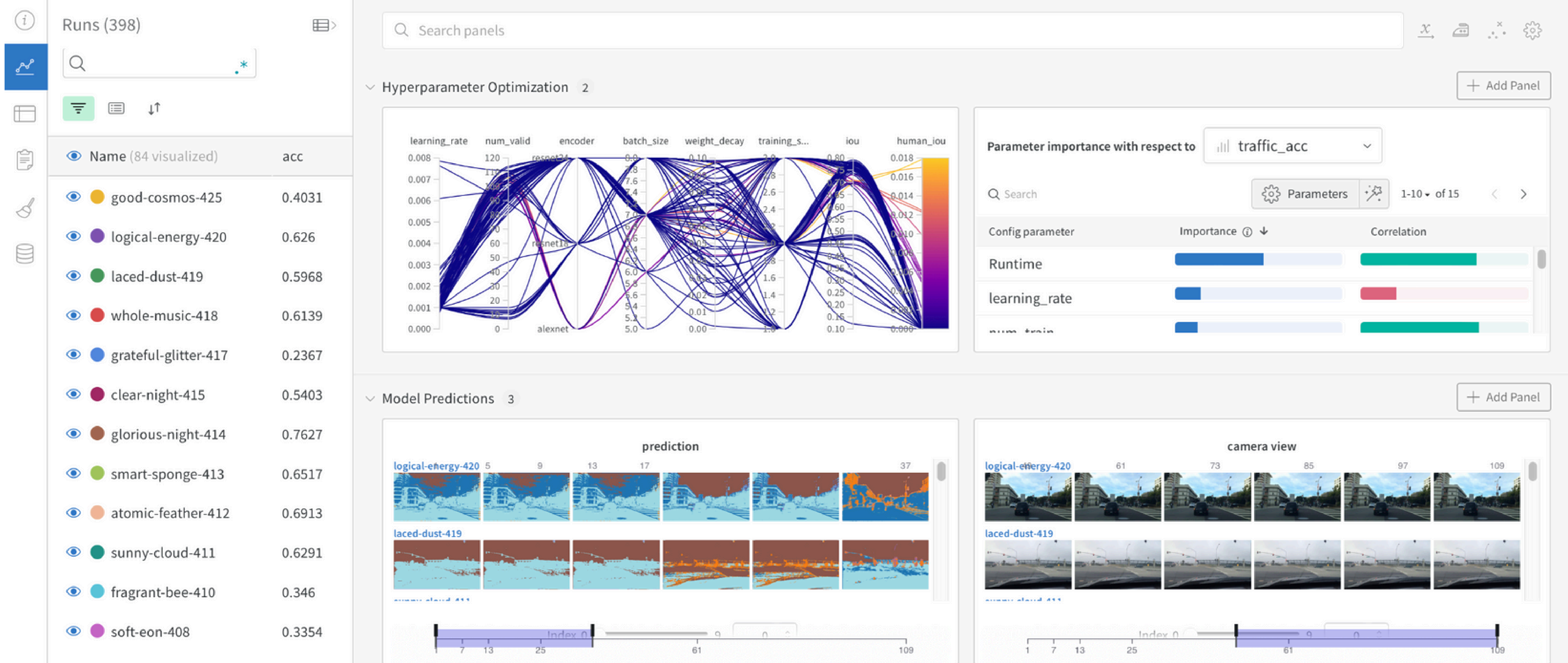
🤝 Collaborate

Share experiments with teammates easily

🔧 Integrates Seamlessly

Works with PyTorch, TensorFlow, scikit-learn, and more

WANDB DASHBOARD



Runs (398)

Name (84 visualized)	acc
good-cosmos-425	0.4031
logical-energy-420	0.626
laced-dust-419	0.5968
whole-music-418	0.6139
grateful-glitter-417	0.2367
clear-night-415	0.5403
glorious-night-414	0.7627
smart-sponge-413	0.6517
atomic-feather-412	0.6913
sunny-cloud-411	0.6291
fragrant-bee-410	0.346
soft-eon-408	0.3354

Search panels

Hyperparameter Optimization 2

Add Panel

Parameter importance with respect to traffic_acc

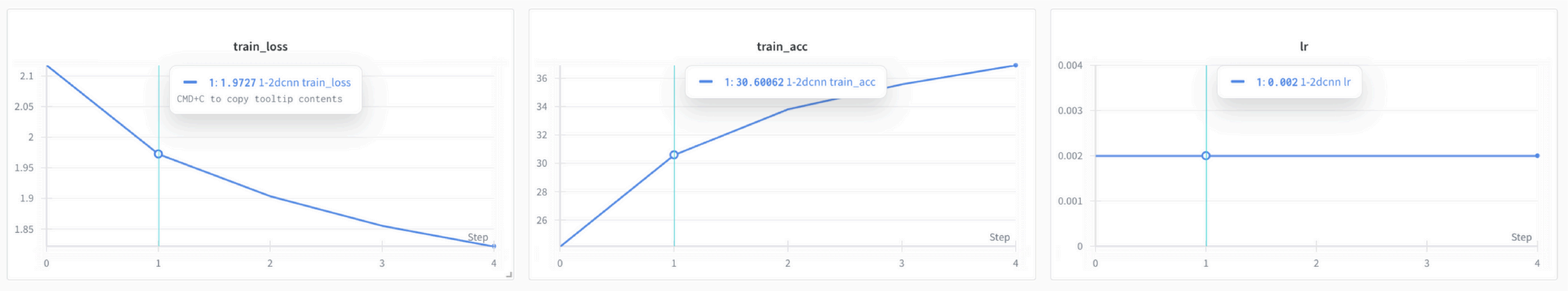
Config parameter	Importance	Correlation
Runtime		
learning_rate		

Model Predictions 3

Add Panel

KEY FEATURES OF WANDB

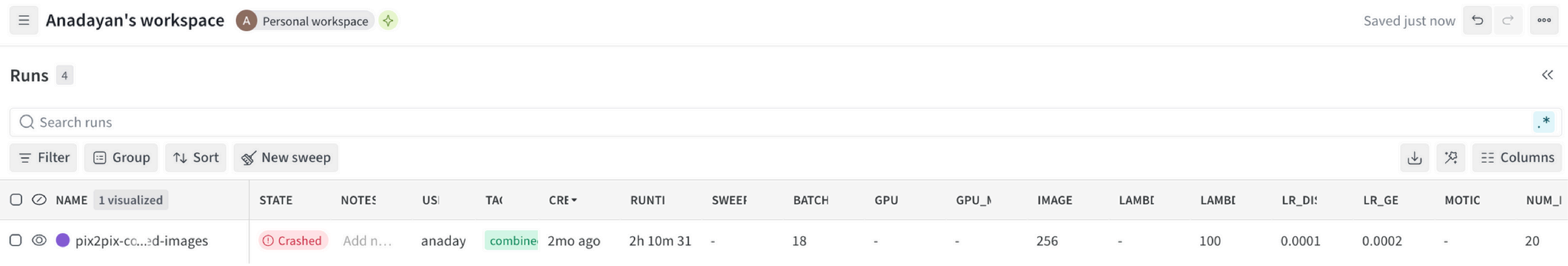
Real-time Performance Logging



- Visualize metrics like Loss and Accuracy F1 in real-time as the model trains.
- Keep track of different models and training runs.
- Spot overfitting or exploding gradients instantly using interactive charts.

KEY FEATURES OF WANDB

Crash Recovery & Resumption



☰ Anadayan's workspace A Personal workspace

Saved just now

Runs 4

Search runs

Filter Group Sort New sweep Columns

□ NAME 1 visualized	STATE	NOTES	USI	TAC	CRE▾	RUNTI	SWEEF	BATCH	GPU	GPU_N	IMAGE	LAMBI	LAMBI	LR_DIS	LR_GE	MOTIC	NUM_I
□ ● pix2pix-co...ed-images	Crashed	Add n...	anaday	combine	2mo ago	2h 10m 31	-	18	-	-	256	-	100	0.0001	0.0002	-	20

- WandB can save your model's checkpoints to the cloud.
- By passing the specific Run ID, you can download the weights and resume the epoch exactly where you left off.

KEY FEATURES OF WANDB

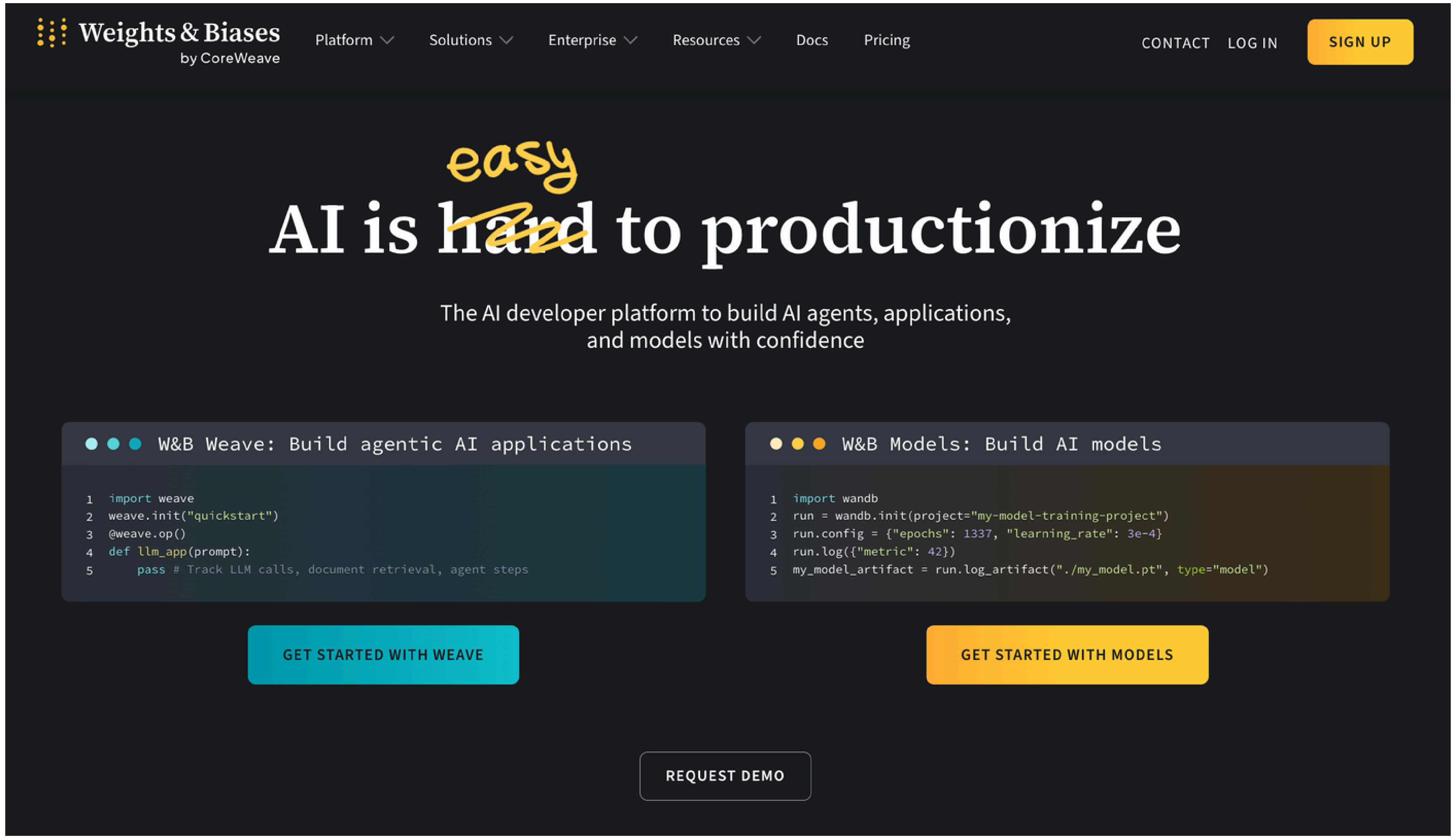
Hyperparameter Tuning

- Automating the search to find the best configuration (Learning Rate, Batch Size, Optimizer).
- Let WandB try different configurations while you sleep!

Collaboration

- All team members log runs to the same "Project" bucket.
- See everyone's runs side-by-side to determine whose model is performing best.

SET UP WANDB ACCOUNT



AI is ~~hard~~ easy to productionize

The AI developer platform to build AI agents, applications, and models with confidence

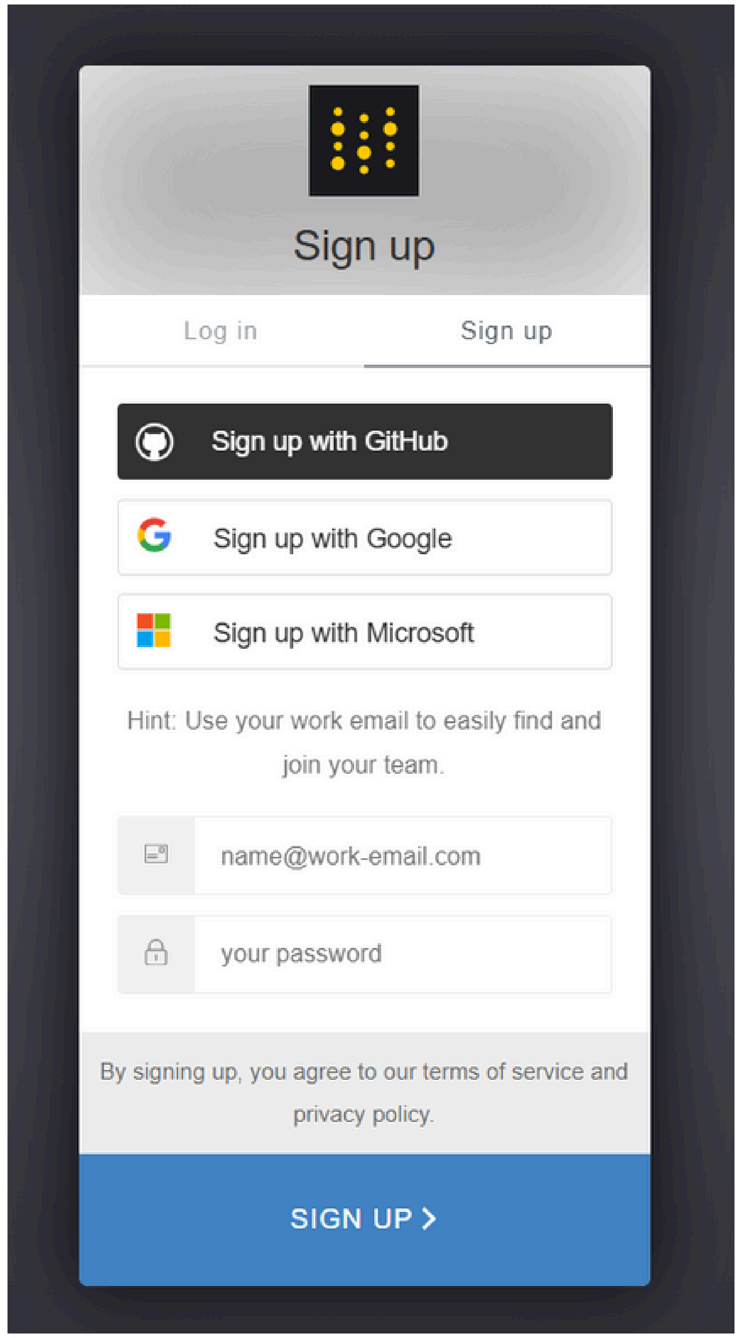
W&B Weave: Build agentic AI applications

```python
import weave
weave.init("quickstart")
@weave.op()
def llm_app(prompt):
    pass # Track LLM calls, document retrieval, agent steps
```

W&B Models: Build AI models

```python
import wandb
run = wandb.init(project="my-model-training-project")
run.config = {"epochs": 1337, "learning_rate": 3e-4}
run.log({"metric": 42})
my_model_artifact = run.log_artifact("./my_model.pt", type="model")
```

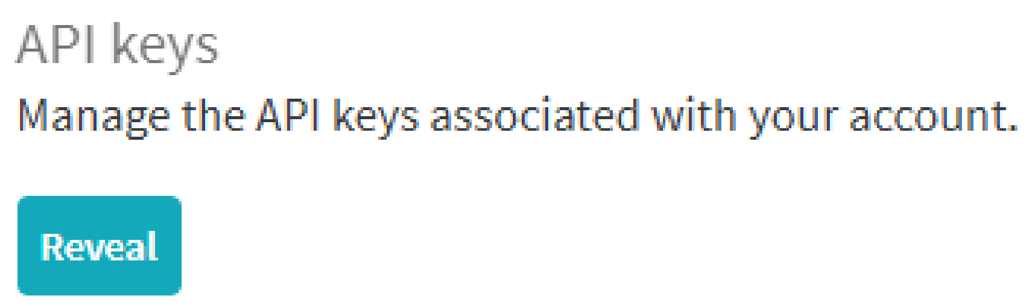
https://wandb.ai/site/



Hint: Use your work email to easily find and join your team.

By signing up, you agree to our terms of service and privacy policy.

SET UP WANDB API KEY



API keys

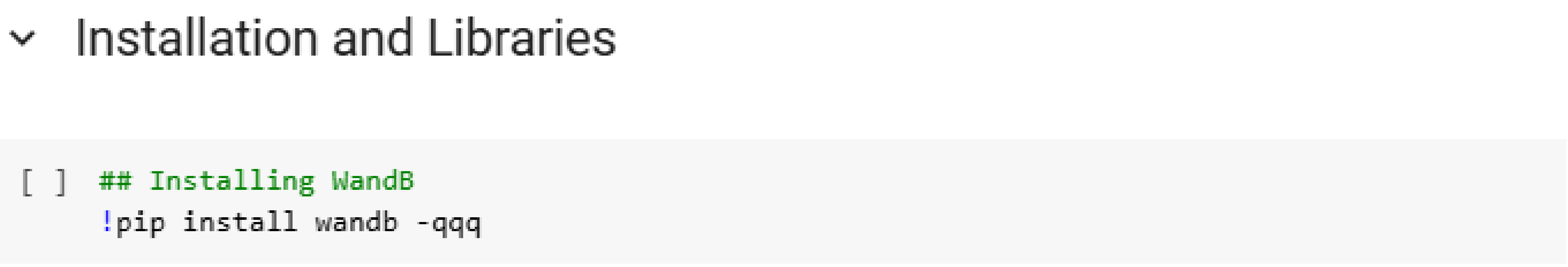
Manage the API keys associated with your account.

- Your API key authenticates your machine/notebook so runs can upload to your WandB dashboard.
- You can find your key in the “API Keys” section in your WandB user settings.

Never post your API key in GitHub, Piazza, screenshots, or shared notebooks!
→ If it does get leaked, please rotate the key immediately in WandB settings.

INSTALLATION & NOTEBOOK SET UP

Wandb Installation for Python



Installation and Libraries

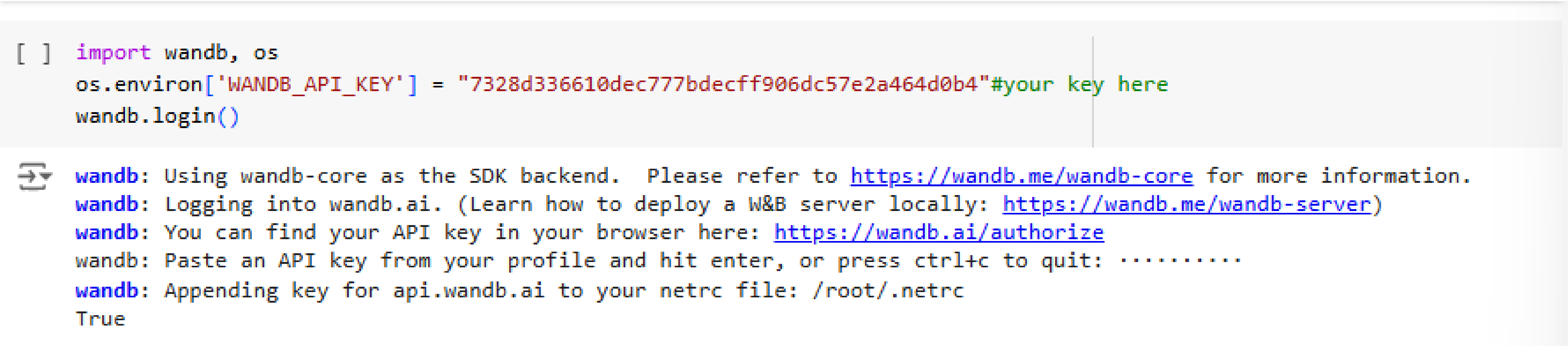
```
## Installing WandB
!pip install wandb -qqq
```

Notebook Setup

- Dataset: CIFAR10
- Neural Network: Convolutional Neural Network
- All functionalities are built in functional blocks for automated access.

Credits to Fall 2024 TAs for the Notebook!

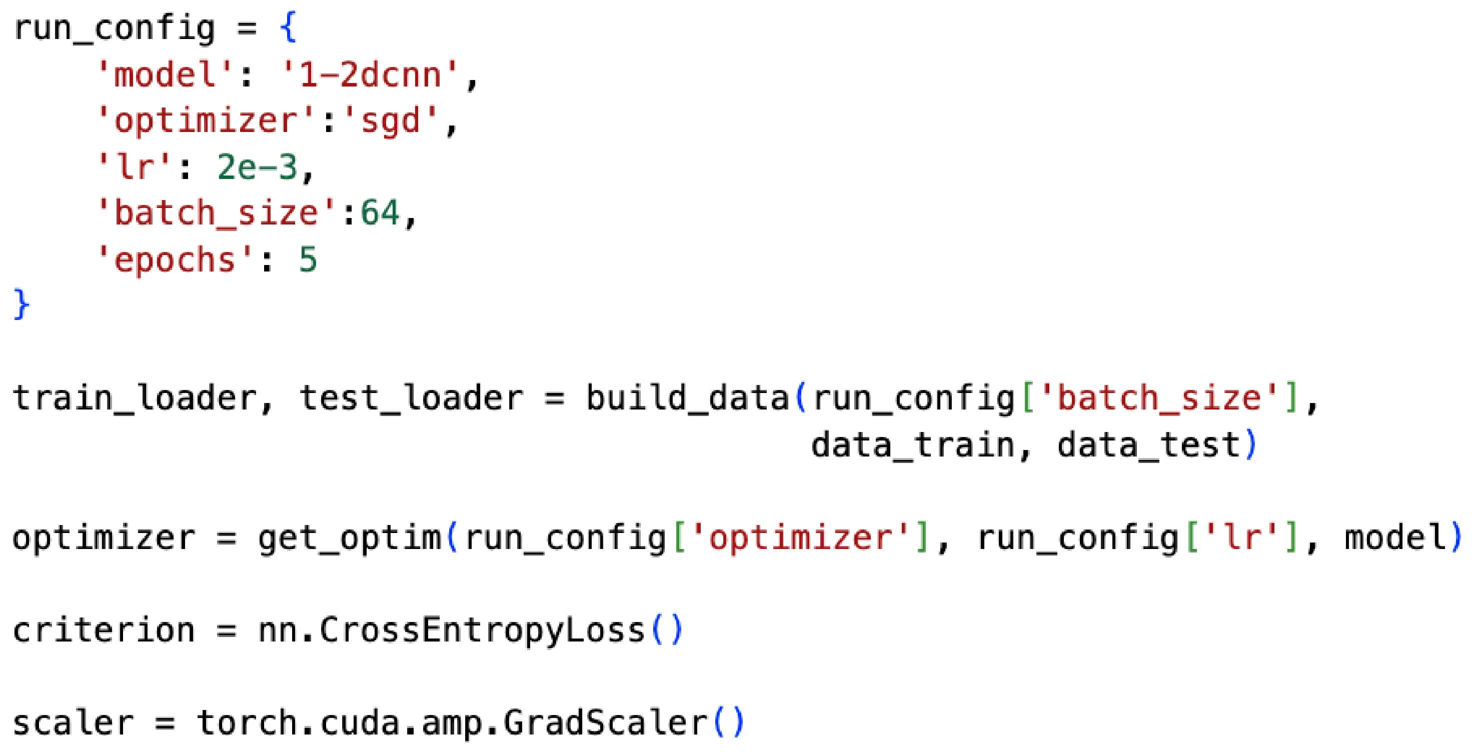
NOTEBOOK: WANDB LOGIN



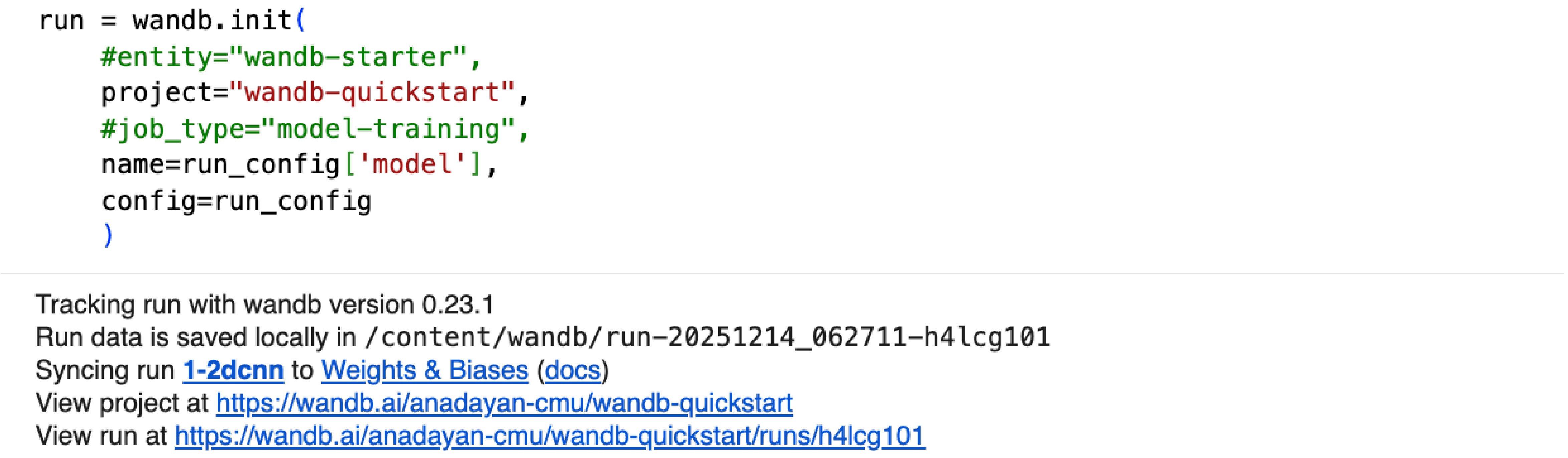
```
[ ] import wandb, os
    os.environ['WANDB_API_KEY'] = "7328d336610dec777bdecff906dc57e2a464d0b4"#your key here
    wandb.login()
```

```
wandb: Using wandb-core as the SDK backend.  Please refer to https://wandb.me/wandb-core for more information.
wandb: Logging into wandb.ai. (Learn how to deploy a W&B server locally: https://wandb.me/wandb-server)
wandb: You can find your API key in your browser here: https://wandb.ai/authorize
wandb: Paste an API key from your profile and hit enter, or press ctrl+c to quit: ··········
wandb: Appending key for api.wandb.ai to your netrc file: /root/.netrc
True
```

NOTEBOOK: INITIALIZING PROJECT



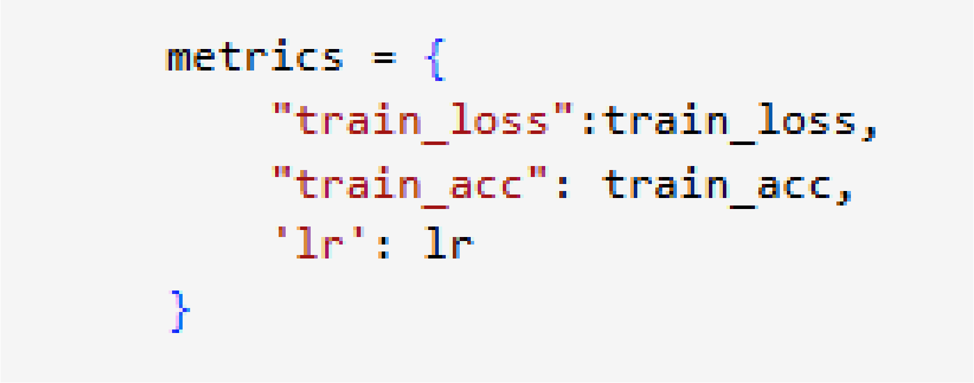
```
run_config = {
    'model': '1-2dcnn',
    'optimizer':'sgd',
    'lr': 2e-3,
    'batch_size':64,
    'epochs': 5
}

train_loader, test_loader = build_data(run_config['batch_size'],
                                       data_train, data_test)

optimizer = get_optim(run_config['optimizer'], run_config['lr'], model)

criterion = nn.CrossEntropyLoss()

scaler = torch.cuda.amp.GradScaler()
```



```
run = wandb.init(
    #entity="wandb-starter",
    project="wandb-quickstart",
    #job_type="model-training",
    name=run_config['model'],
    config=run_config
    )
```

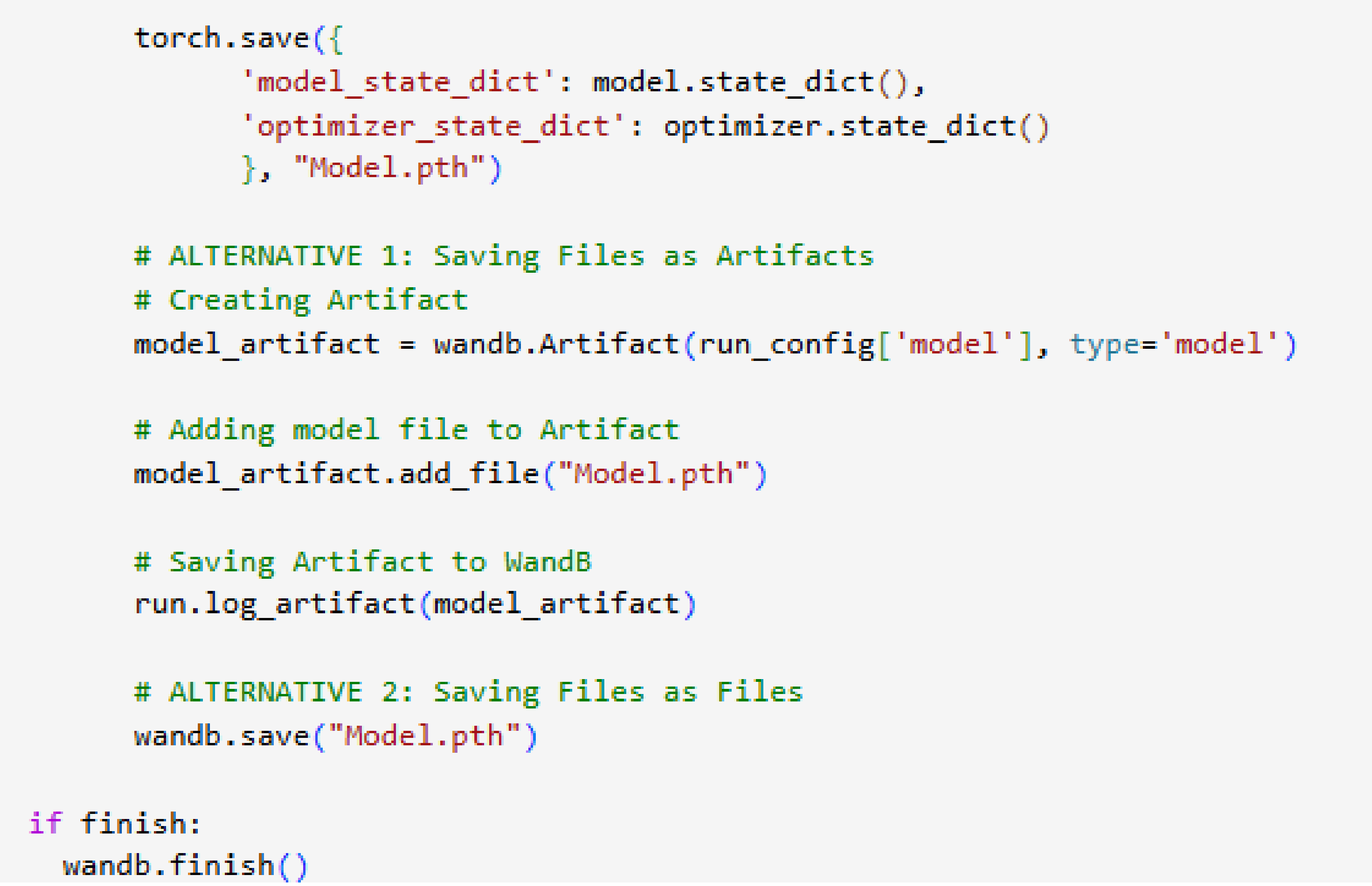
```
Tracking run with wandb version 0.23.1
Run data is saved locally in /content/wandb/run-20251214_062711-h4lcg101
Syncing run 1-2dcnn to Weights & Biases (docs)
View project at https://wandb.ai/anadayan-cmu/wandb-quickstart
View run at https://wandb.ai/anadayan-cmu/wandb-quickstart/runs/h4lcg101
```

NOTEBOOK: LOGGING & SAVING MODELS



```
metrics = {
    "train_loss":train_loss,
    "train_acc": train_acc,
    'lr': lr
}
```

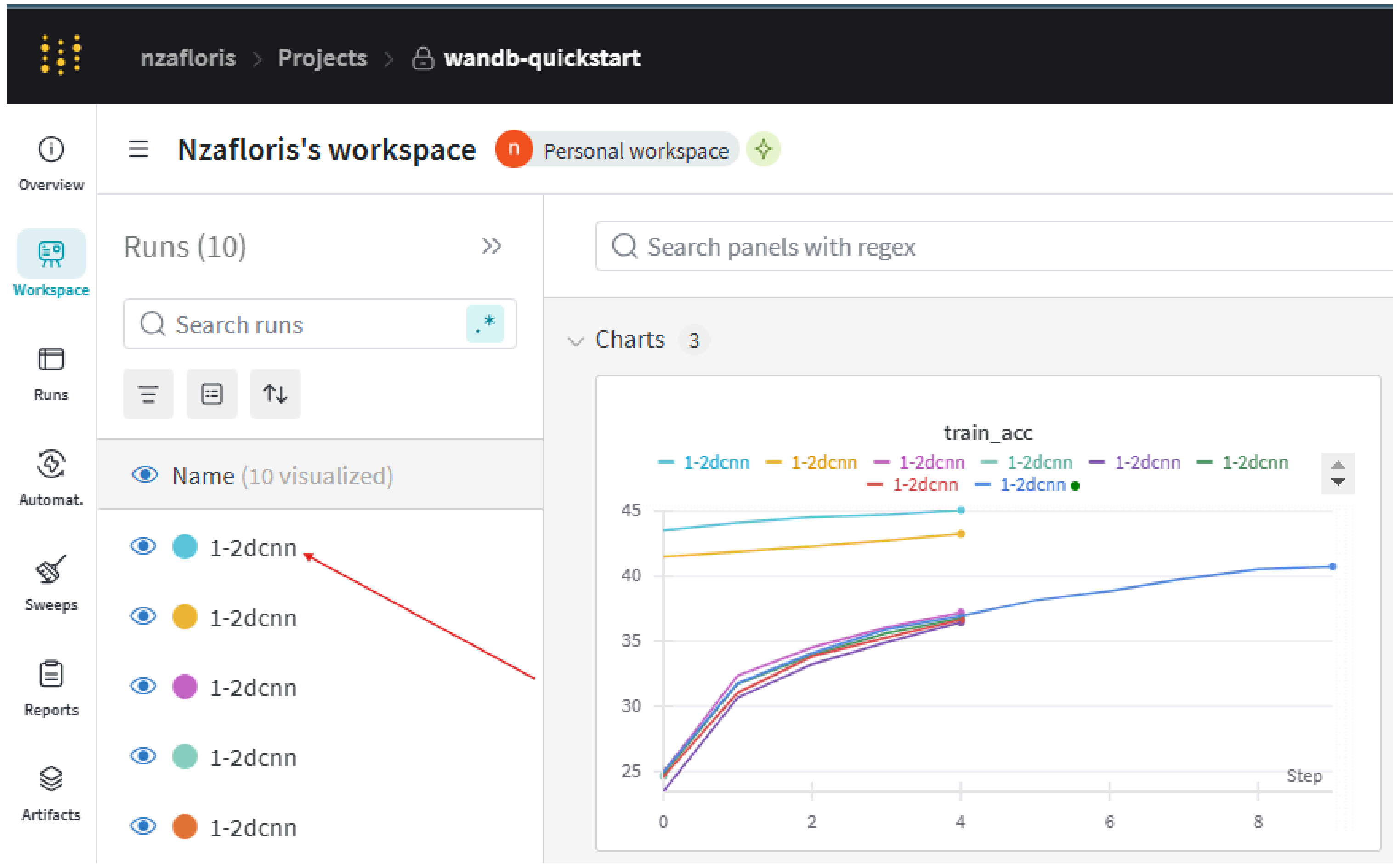
Logging metrics

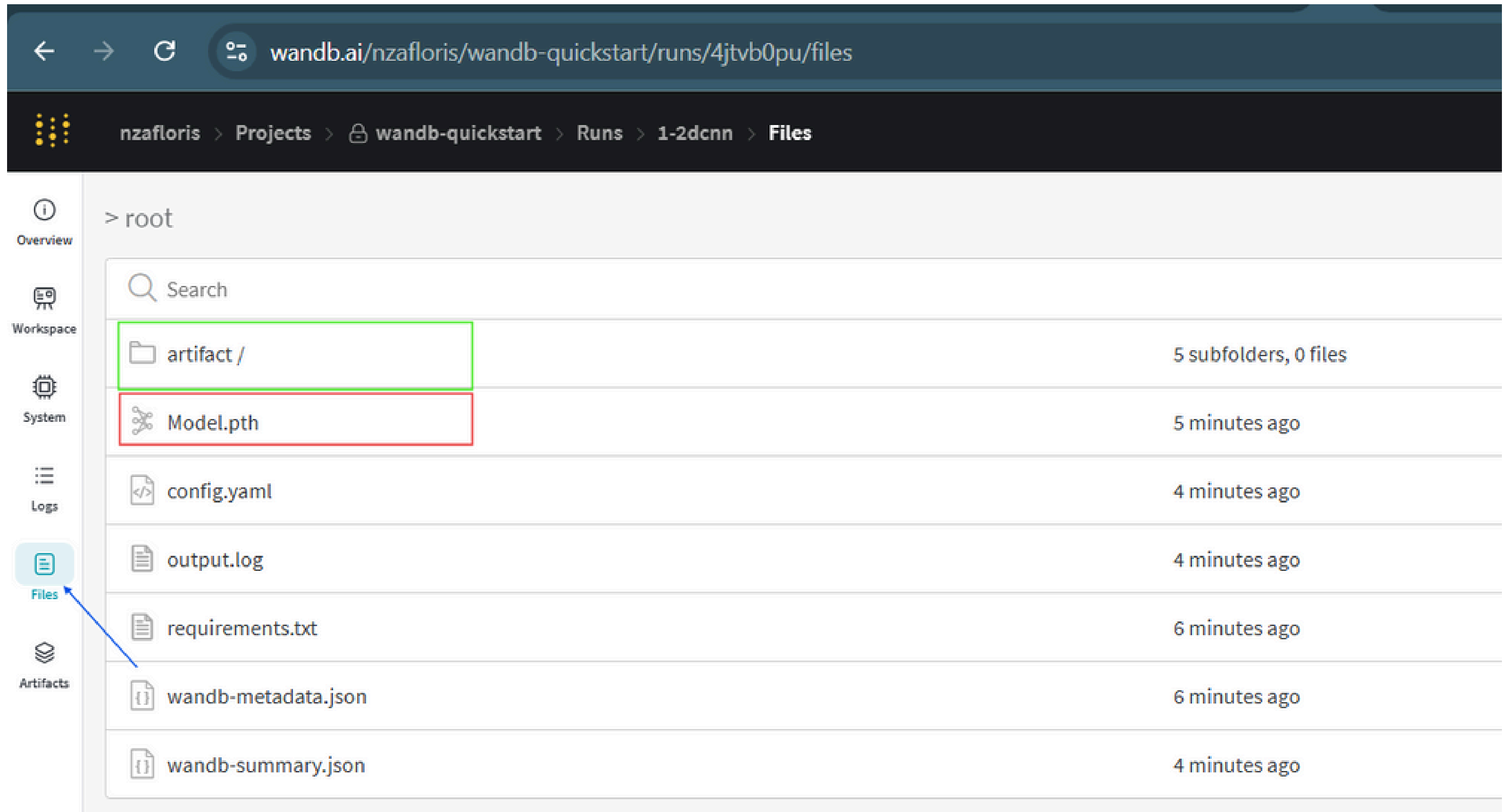


```
      torch.save({
            'model_state_dict': model.state_dict(),
            'optimizer_state_dict': optimizer.state_dict()
            }, "Model.pth")

      # ALTERNATIVE 1: Saving Files as Artifacts
      # Creating Artifact
      model_artifact = wandb.Artifact(run_config['model'], type='model')

      # Adding model file to Artifact
      model_artifact.add_file("Model.pth")

      # Saving Artifact to WandB
      run.log_artifact(model_artifact)

      # ALTERNATIVE 2: Saving Files as Files
      wandb.save("Model.pth")

if finish:
  wandb.finish()
```

Saving models

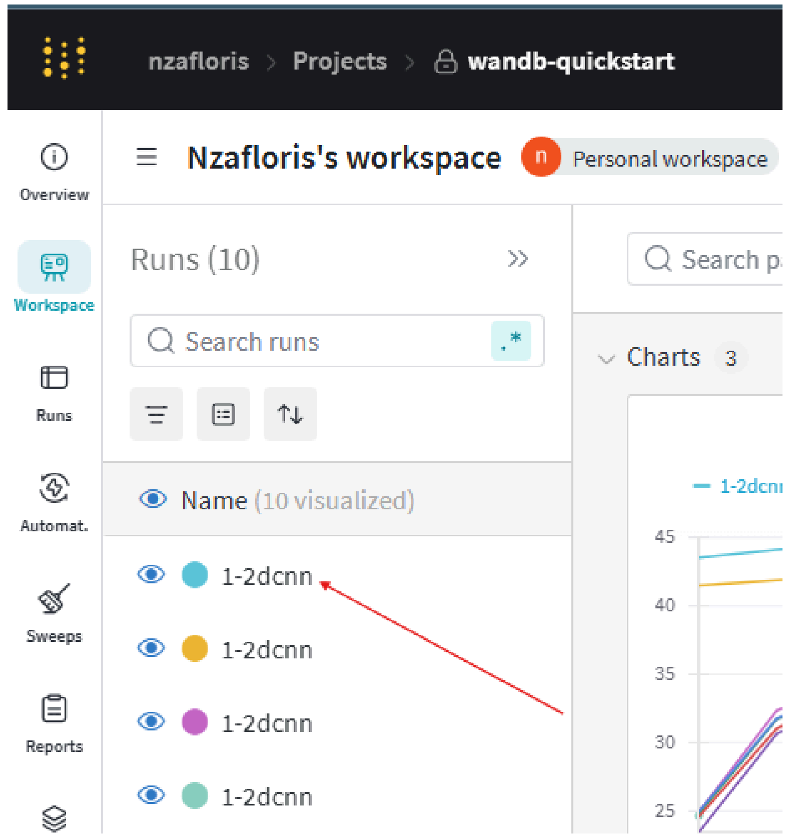
NOTEBOOK: LOGGING & SAVING MODELS

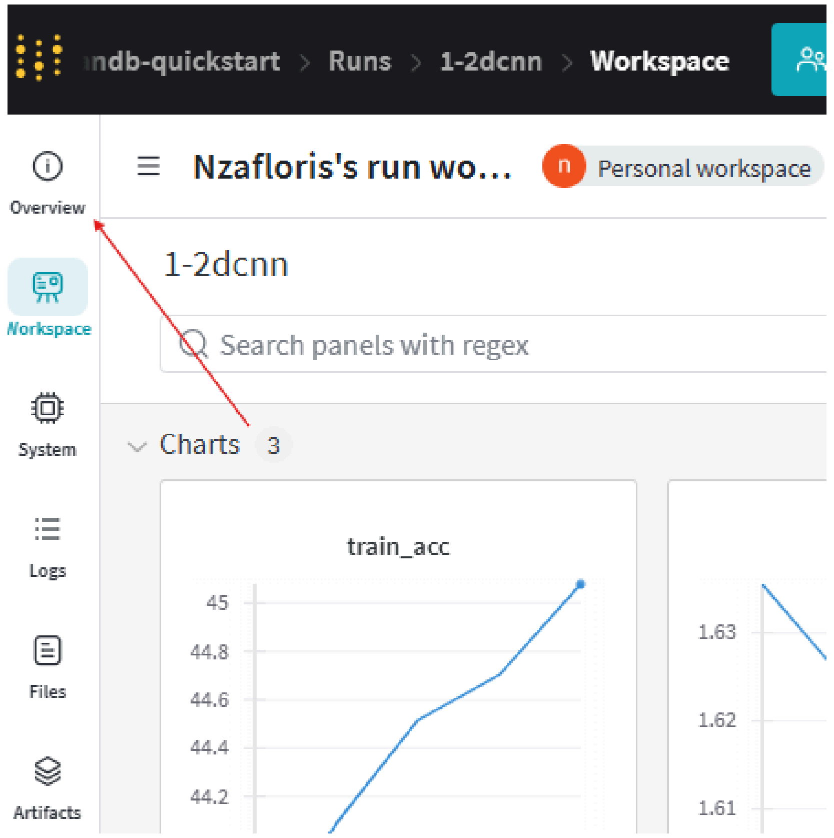


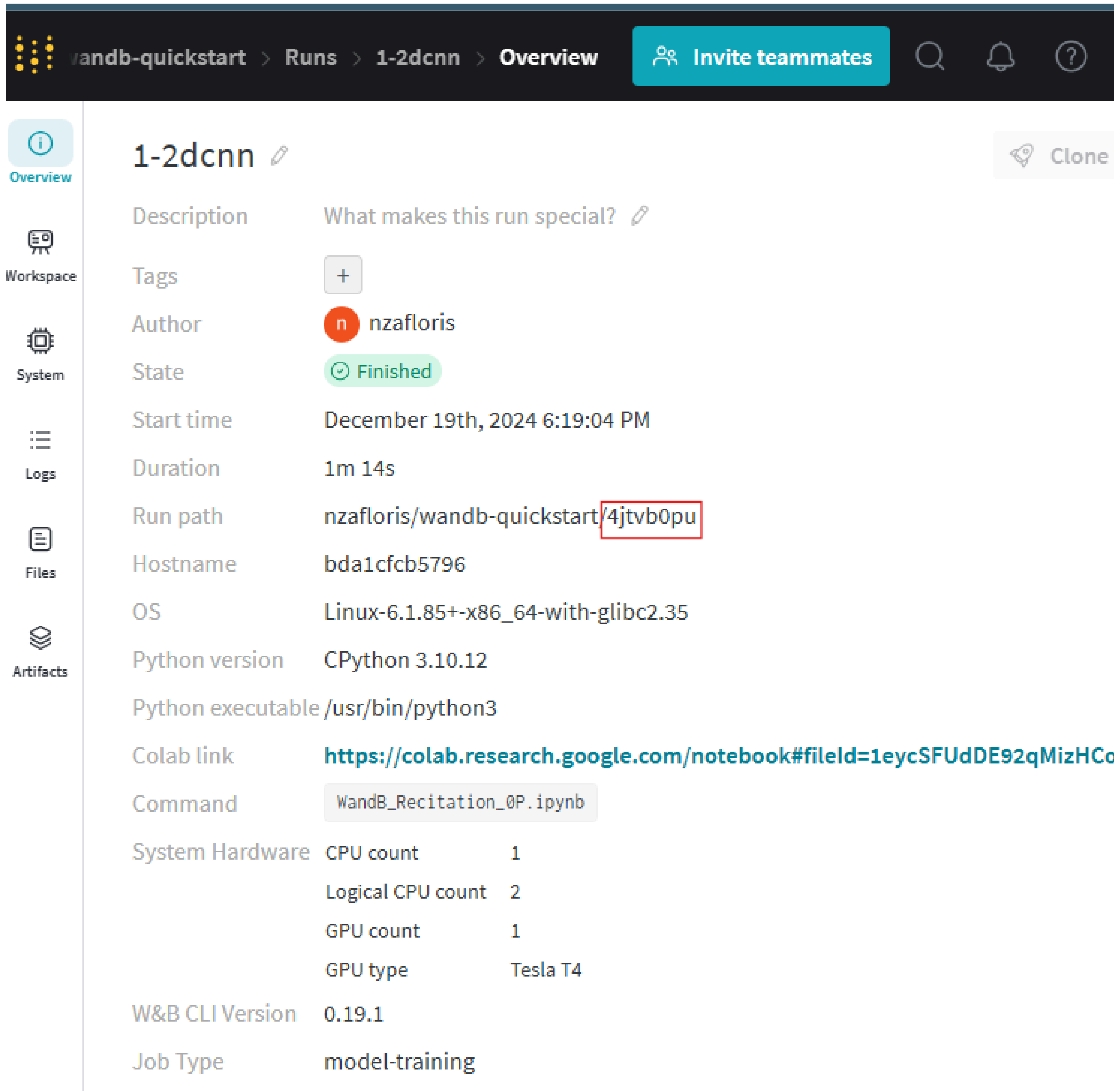

wandb.ai/nzafloris/wandb-quickstart/runs/4jtvb0pu/files
nzafloris > Projects > wandb-quickstart > Runs > 1-2dcnn > Files
Overview
Workspace
System
Logs
Files
Artifacts
> root
Search

artifact /	5 subfolders, 0 files
Model.pth	5 minutes ago
config.yaml	4 minutes ago
output.log	4 minutes ago
requirements.txt	6 minutes ago
wandb-metadata.json	6 minutes ago
wandb-summary.json	4 minutes ago

NOTEBOOK: GET RUN ID







1-2dcnn

Description What makes this run special?

Tags +

Author nzafloris

State Finished

Start time December 19th, 2024 6:19:04 PM

Duration 1m 14s

Run path nzafloris/wandb-quickstart/4jtvb0pu

Hostname bda1cfcb5796

OS Linux-6.1.85+-x86_64-with-glibc2.35

Python version CPython 3.10.12

Python executable /usr/bin/python3

Colab link https://colab.research.google.com/notebook#fileId=1eycSFUdDE92qMizHCo

Command `WandB_Recitation_0P.ipynb`

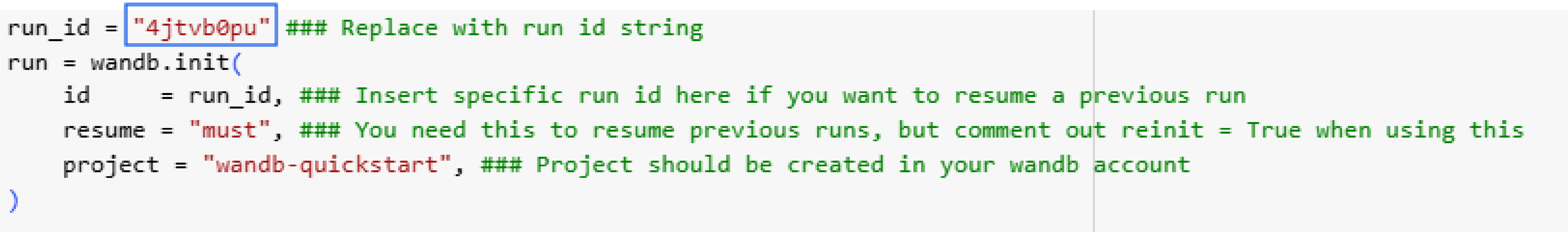
System Hardware

CPU count	1
Logical CPU count	2
GPU count	1
GPU type	Tesla T4

W&B CLI Version 0.19.1

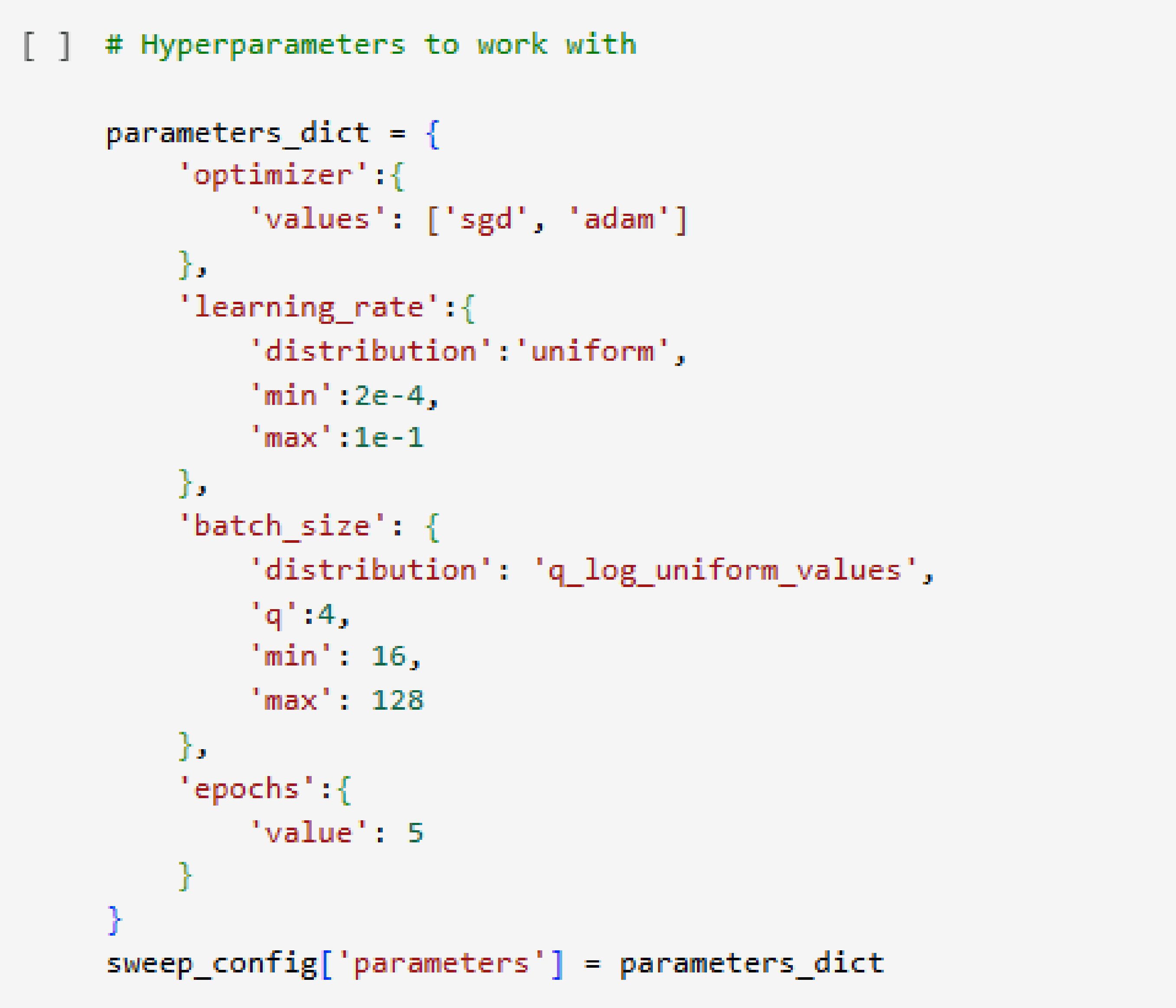
Job Type model-training

NOTEBOOK: RESUMING RUN

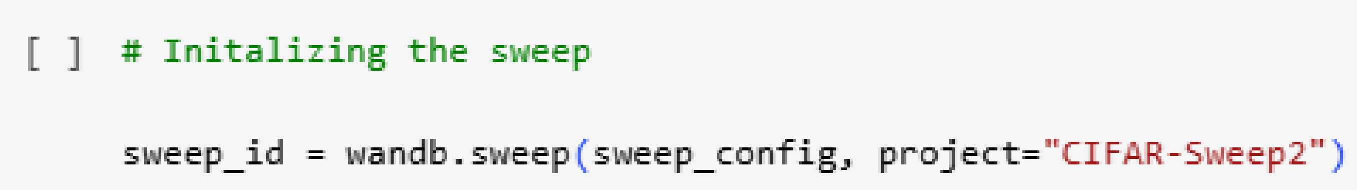


```
run_id = "4jtvb0pu" ### Replace with run id string
run = wandb.init(
    id     = run_id, ### Insert specific run id here if you want to resume a previous run
    resume = "must", ### You need this to resume previous runs, but comment out reinit = True when using this
    project = "wandb-quickstart", ### Project should be created in your wandb account
)
```

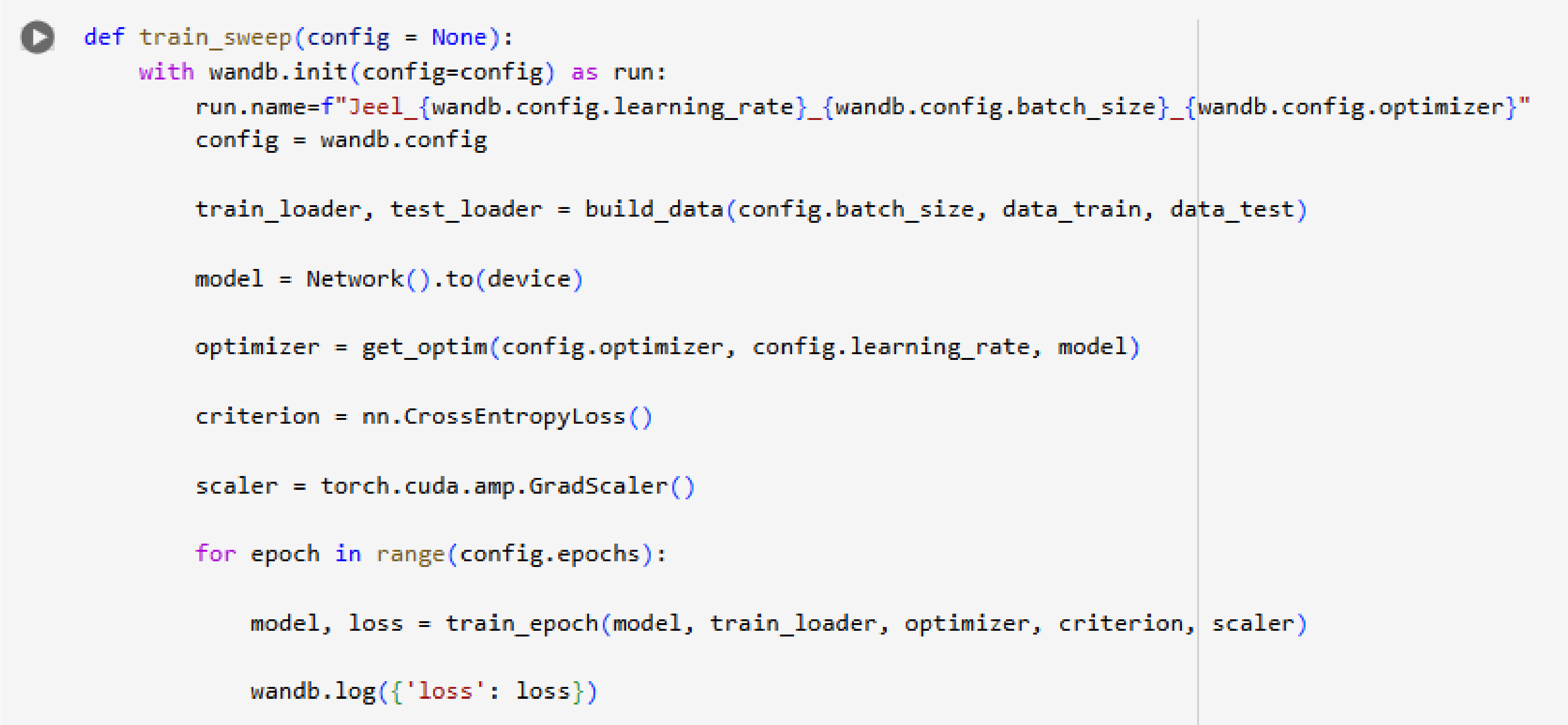
NOTEBOOK: HYPERPARAMETER SWEEPING



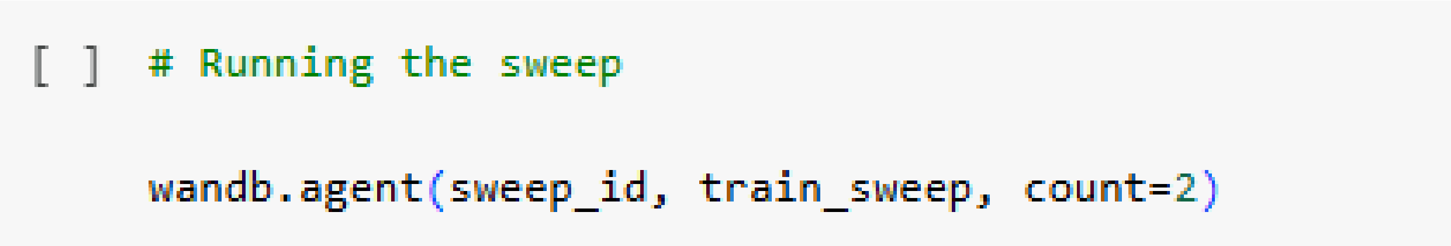
```
[ ]  # Hyperparameters to work with

     parameters_dict = {
         'optimizer':{
             'values': ['sgd', 'adam']
         },
         'learning_rate':{
             'distribution':'uniform',
             'min':2e-4,
             'max':1e-1
         },
         'batch_size': {
             'distribution': 'q_log_uniform_values',
             'q':4,
             'min': 16,
             'max': 128
         },
         'epochs':{
             'value': 5
         }
     }
     sweep_config['parameters'] = parameters_dict
```



```
[ ]  # Initalizing the sweep

     sweep_id = wandb.sweep(sweep_config, project="CIFAR-Sweep2")
```

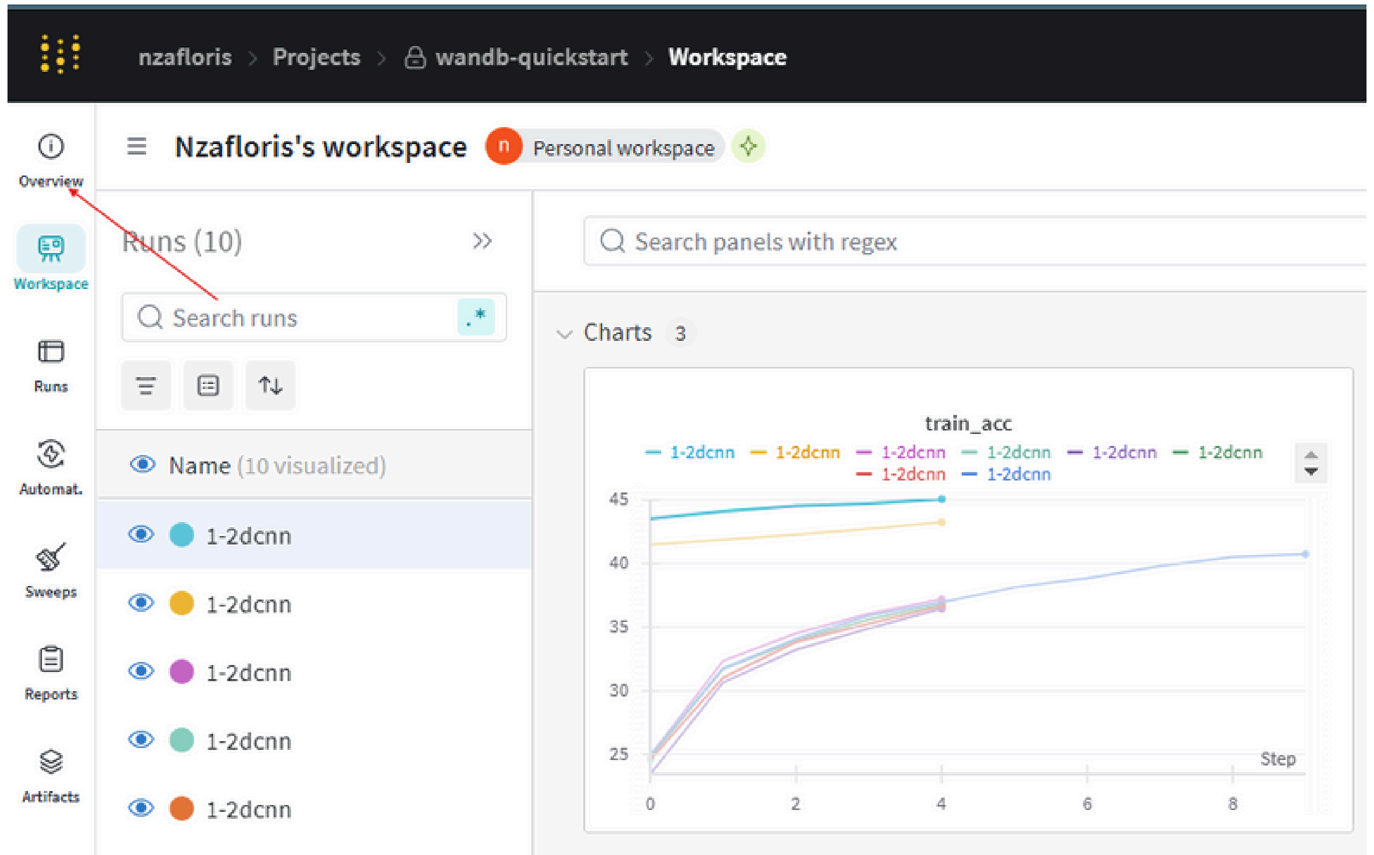


```
def train_sweep(config = None):
    with wandb.init(config=config) as run:
        run.name=f"Jeel_{wandb.config.learning_rate}_{wandb.config.batch_size}_{wandb.config.optimizer}"
        config = wandb.config

        train_loader, test_loader = build_data(config.batch_size, data_train, data_test)

        model = Network().to(device)

        optimizer = get_optim(config.optimizer, config.learning_rate, model)

        criterion = nn.CrossEntropyLoss()

        scaler = torch.cuda.amp.GradScaler()

        for epoch in range(config.epochs):

            model, loss = train_epoch(model, train_loader, optimizer, criterion, scaler)

            wandb.log({'loss': loss})
```



```
[ ]  # Running the sweep

     wandb.agent(sweep_id, train_sweep, count=2)
```

MAKING YOUR WANDB PROJECT PUBLIC



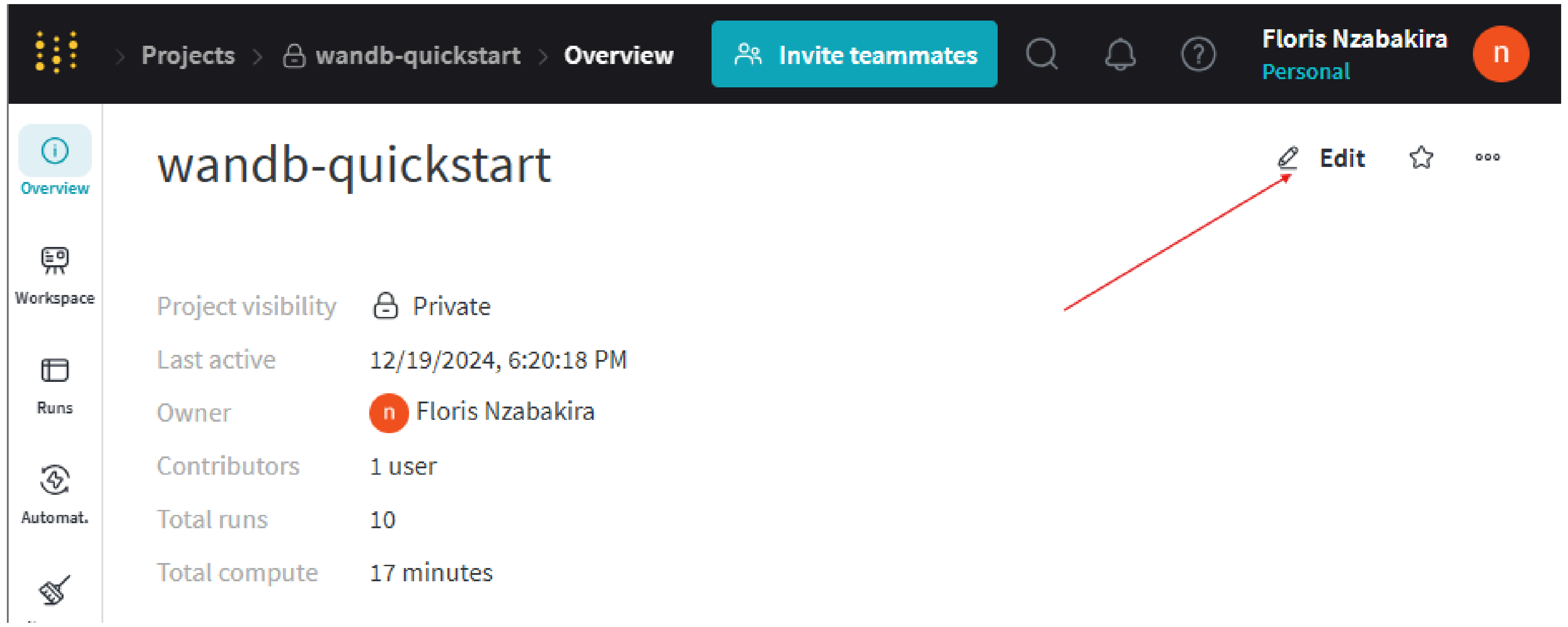
Nzafloris's workspace Personal workspace

Runs (10)

Name (10 visualized)

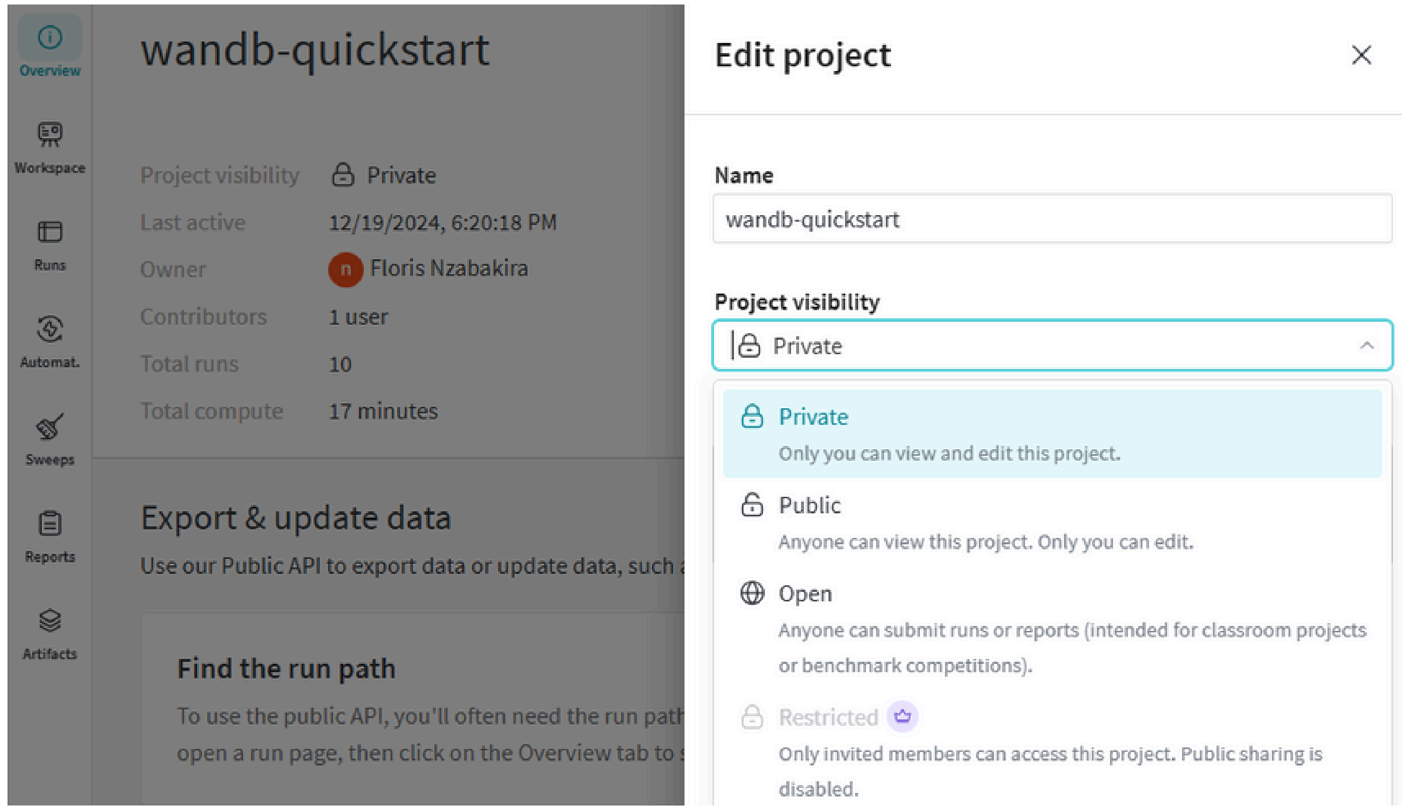
- 1-2dcnn
- 1-2dcnn
- 1-2dcnn
- 1-2dcnn
- 1-2dcnn

Charts 3



wandb-quickstart

Project visibility	Private
Last active	12/19/2024, 6:20:18 PM
Owner	Floris Nzabakira
Contributors	1 user
Total runs	10
Total compute	17 minutes



wandb-quickstart

Project visibility	Private
Last active	12/19/2024, 6:20:18 PM
Owner	Floris Nzabakira
Contributors	1 user
Total runs	10
Total compute	17 minutes

Export & update data

Use our Public API to export data or update data, such

Find the run path

To use the public API, you'll often need the run path

open a run page, then click on the Overview tab to

Edit project

Name

wandb-quickstart

Project visibility

Private

- Private
  Only you can view and edit this project.
- Public
  Anyone can view this project. Only you can edit.
- Open
  Anyone can submit runs or reports (intended for classroom projects or benchmark competitions).
- Restricted
  Only invited members can access this project. Public sharing is disabled.

WANDB STORAGE FOR ACCOUNTS

Once your WandB free storage is full, you won't be able to access your runs!

- Individual Account:
  - Provides 100GB free
  - Can be accessed through https://wandb.ai/subscriptions

- Group Account:
  - Provides 5GB free per team
  - Can be accessed through:
    https://wandb.ai/account-settings/<Team-Name>/billing

THANK YOU!