

# MLOps

Part 5.

## Automate Hyperparameter Tuning with W&B Sweeps



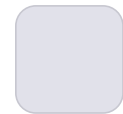
POLITÉCNICA

UNIVERSIDAD  
POLITÉCNICA  
DE MADRID



Máster  
Deep Learning

# Why Hyperparameter Optimization Matters

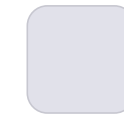


Hyperparameters significantly impact model performance

Examples include learning rate, batch size, and layer size



Manual tuning is tedious, time-consuming, and often suboptimal



W&B Sweeps automates this process

## What Sweeps Do:

- Define a **search space** for your hyperparameters
- Specify a **search strategy** (random, grid, bayes)
- Coordinate multiple training **runs** (executed by agents)
- Track results and help you find the **best** combination

# The 4 Key Steps to Running a Sweep

Set up Training Code

Integrate W&B (wandb.init,  
wandb.config, wandb.log)

Define Search Space  
Create a Sweep Configuration  
(YAML or Python Dict)

Initialize Sweep  
Tell the W&B server about  
your sweep (get a sweep\_id)

Start Agent(s)  
Launch workers to run the  
experiments

# Instrumenting Your Training Script

## Key Components

- **wandb.init():** Initialize a W&B run within your training function. *Crucial:* Ensure the project matches your sweep project
- **wandb.config:** Access hyperparameters provided by the Sweep Agent. Replace hardcoded values
- **wandb.log({"metric\_name": value}):** Log the metric you want to optimize (e.g., val\_loss, accuracy). *Crucial:* The logged metric name *must exactly match* the name in your sweep configuration

## Code Snippet

```
import wandb

def main():
    wandb.init(project="my-sweep-project") # Project
    name matters!

    # Access sweep-provided hparams
    lr = wandb.config.learning_rate

    # ... train model ...
    accuracy = # ... calculate metric ...

    wandb.log({"accuracy": accuracy}) # Log the
    metric to optimize
```

# Specifying the Search Space



## Sweep Configuration

Defines *what* to search and *how*. Can be a YAML file or Python dictionary.

## Key Components

- **method:** Search strategy (random, grid, bayes)
- **metric:** What to optimize (name of the logged metric, goal: minimize or maximize)
- **parameters:** The hyperparameters to explore

Simple Example (YAML):

```
method: random
metric:
  name: val_loss
  goal: minimize
parameters:
  learning_rate:
    min: 0.001
    max: 0.1
  batch_size:
    values: [16, 32, 64]
```

# Exploring Hyperparameter Options



## Common Parameter Types

- **values:** List of discrete options (e.g., [16, 32, 64])
- **Ranges:** min, max (for continuous values like learning rate)
- **Distributions:** uniform, log\_uniform, etc. (often used with bayes or random)

## Search Methods

- **grid:** Tries every combination (can be slow)
- **random:** Samples random combinations (good baseline)
- **bayes:** Uses previous results to pick promising next parameters (often most efficient)

**Early Stopping:** You can use `early_terminate` (e.g., hyperband) for stopping unpromising runs early.

# Starting the Sweep Controller



## What Happens

Sends your configuration to the W&B server (or runs locally). Creates a "Sweep Controller" that manages the process. Returns a unique sweep\_id (looks like entity/project/sweep\_id).



## Python Initialization

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



## CLI Initialization

```
wandb sweep config.yaml --project my-sweep-project
```

**Crucial:** You need this sweep\_id for the next step!

# Launching the Experiment Workers



## Agents Connect

Agents connect to the Sweep Controller using the sweep\_id



## Request Parameters

They ask for a set of hyperparameters



## Run Training

Run your training script (main() function or program specified in config)



## Report Results

Send metrics back to W&B

How to Start Agents:

Python

```
wandb.agent(sweep_id, function=main, count=10)
```

CLI

```
wandb agent entity/project/sweep_id --count 10
```

The count parameter limits the number of runs

**Stopping:** Use Ctrl+C in the terminal, or set count. random/bayes sweeps run indefinitely otherwise.

You can run multiple agents in parallel (on different machines, GPUs, or CPU cores) to speed up the sweep.



# Understanding Your Sweep's Performance

W&B automatically creates visualizations in the UI under the "Sweeps" tab of your project.

## Parallel Coordinates Plot

Shows relationships between all hyperparameters and the target metric. Identify trends.

## Parameter Importance Plot

Ranks hyperparameters by their correlation with the target metric. See what matters most.

## Scatter Plot

Compare runs based on any two parameters or metrics. Spot high-performing runs.

These visualizations help you understand which hyperparameters have the most impact on your model's performance and identify the optimal configuration.

# Recap & Further Exploration

## Summary

- Sweeps automate hyperparameter tuning, saving time and potentially finding better models
- Follow the 4 steps: Instrument Code -> Define Config -> Initialize Sweep -> Start Agent(s)
- Visualize results in the W&B UI to gain insights

## Next Steps / Advanced Topics

- Try different search strategies (bayes)
- Implement early stopping (hyperband)
- Run agents in parallel for faster results
- Explore creating sweeps from the UI or using local controllers

## Resources

- Link to W&B Sweeps Documentation ([docs.wandb.ai/guides/sweeps/](https://docs.wandb.ai/guides/sweeps/))
- Link to relevant examples or the Sweeps GitHub repo