

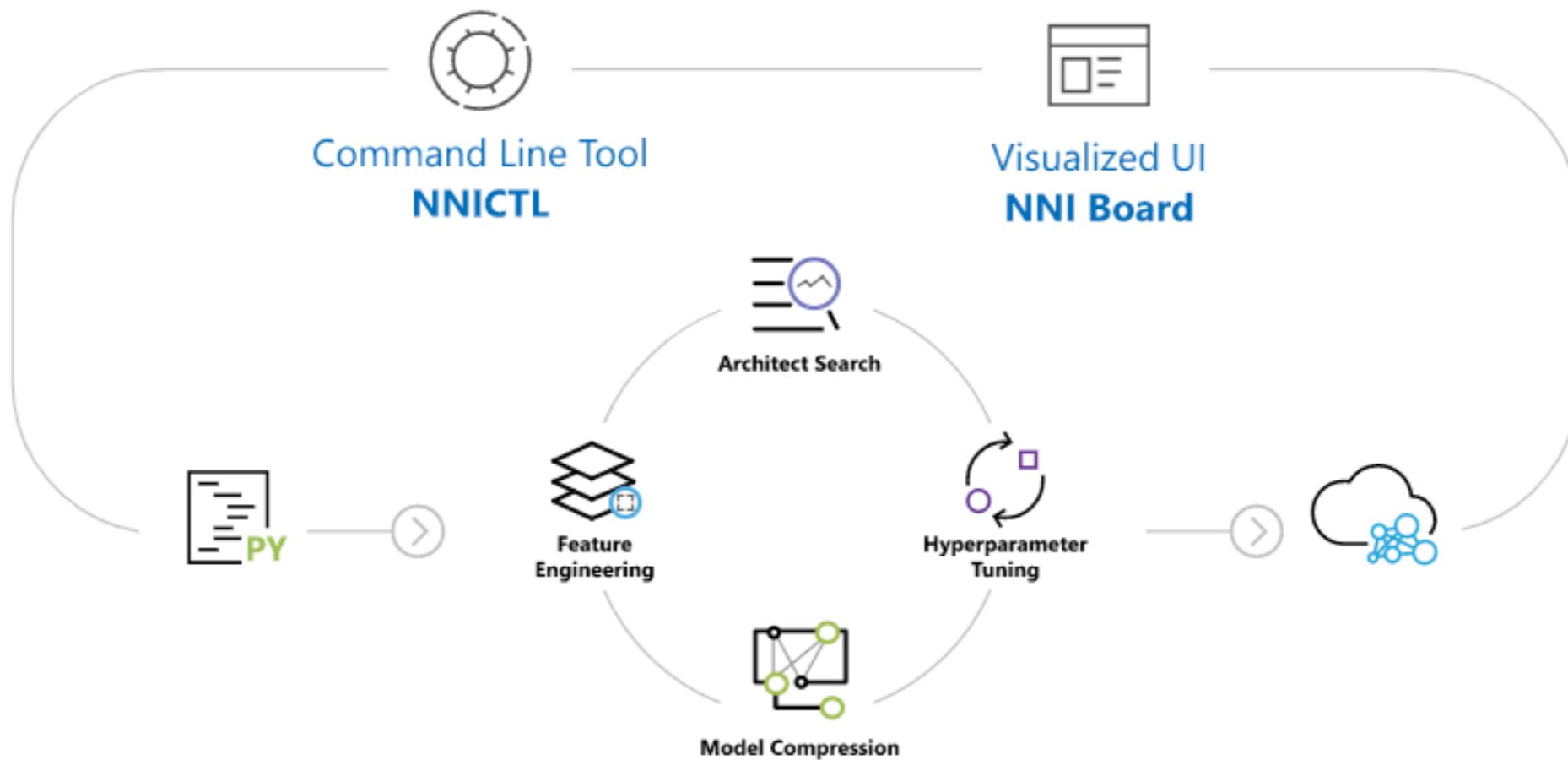
Lab 04 Model Tuning

2021. 08. 07(Sat)
캐글 코리아
강천성

AutoML: NNI

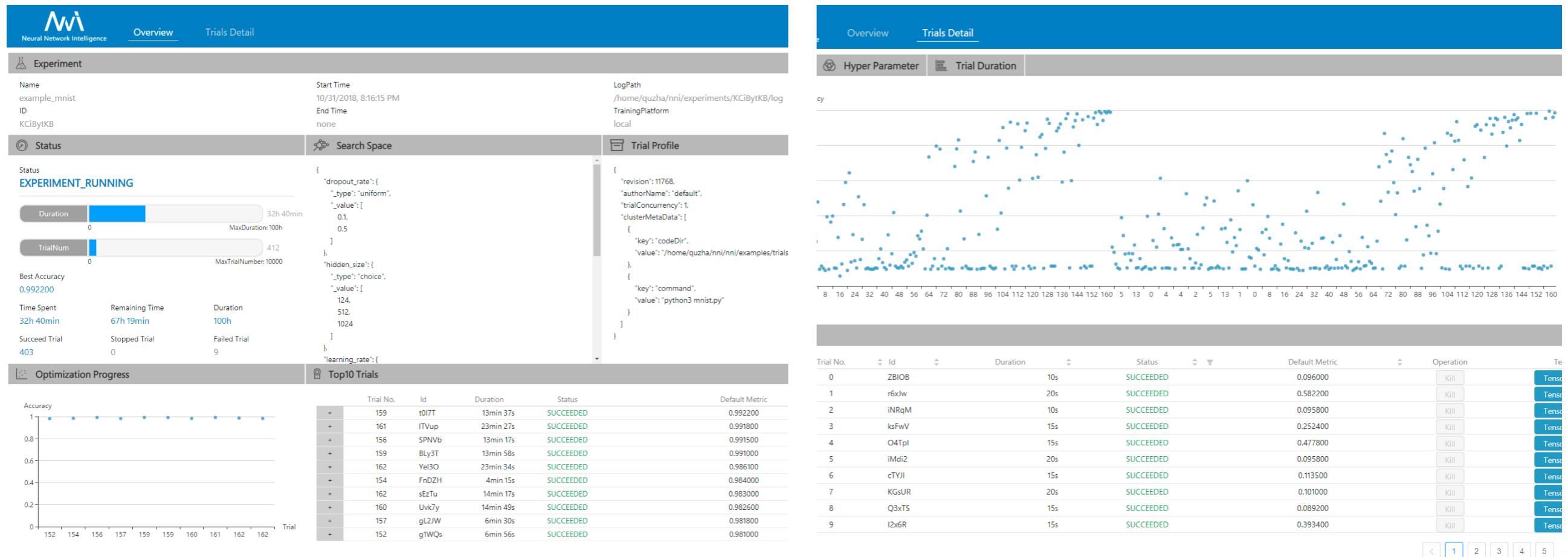
Neural Network Intelligence (Microsoft)

- 제공하는 기능
 - Feature Engineering
 - Neural Architecture Search
 - **Hyperparameter Tuning**
 - Model Compression



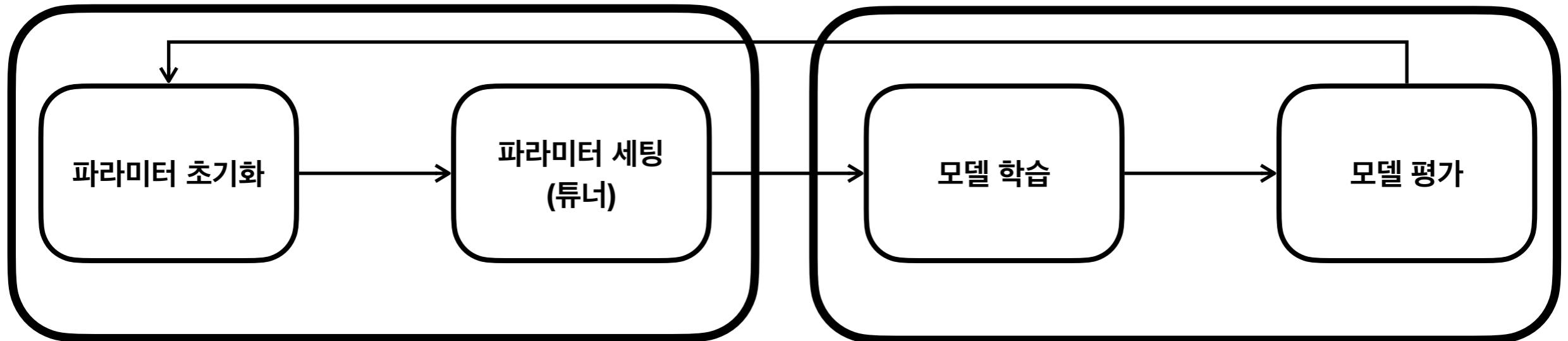
AutoML: NNI

NNI Dashboard



AutoML: NNI

하이퍼파라미터 튜닝 흐름 - trial.py



- 기본 파라미터 세팅
- 파라미터 세팅
- 데이터 전처리
- 모델 학습 및 검증
- 모델 평가
- 결과 생성 등..

AutoML: NNI

Hyperparameter Tuning - config.yml

```
authorName: Cheonseong
experimentName: Lab04
trialConcurrency: 2
maxExecDuration: 48h
maxTrialNum: 99999
#choice: local, remote, pai
trainingServicePlatform: local
#nniManagerIp:
#choice: true, false
searchSpacePath: search_space.json
useAnnotation: false
tuner:
  #choice: TPE, Random, Anneal, Evolution, BatchTuner, MetisTuner
  #SMAC (SMAC should be installed through nnictl)
  builtinTunerName: Anneal
  classArgs:
    #choice: maximize, minimize
    optimize_mode: minimize
trial:
  command: python3 trial.py
  codeDir: .
```

AutoML: NNI

Hyperparameter Tuning - search_space.json

```
{  
    "max_depth": {"_type": "randint", "_value": [2, 15]},  
    "subsample": {"_type": "uniform", "_value": [0.5, 1.0]},  
    "colsample_bytree": {"_type": "uniform", "_value": [0.5, 1.0]}  
}
```

- 임의의 값 선택
 `{"_type": "choice", "_value": ["value1", "value2"]}`
- 임의의 정수 범위
 `{"_type": "randint", "_value": ["lower", "upper"]}`
- 임의의 실수 범위: Uniform 분포에서 추출
 `{"_type": "uniform", "_value": ["low", "high"]}`
- 임의의 실수 범위: Log Uniform 분포에서 추출
 `{"_type": "loguniform", "_value": ["log", "high"]}`
- 임의의 실수 범위: Normal 분포에서 추출
 `{"_type": "normal", "_value": ["mu(mean)", "sigma(standard deviation)"]}`
- 임의의 실수 범위: Log Normal 분포에서 추출
 `{"_type": "lognormal", "_value": ["mu(mean)", "sigma(standard deviation)"]}`

AutoML: NNI

설치 및 사용법

- Terminal 또는 Anaconda Powershell Prompt에서 “pip install xxx” 명령어로 nni 설치
 - Windows 환경 시 C++ 14.0 버전을 먼저 설치해야 합니다.
 - 다음 링크를 따라 동일하게 설치합니다. 이미 설치 되어있다는 창이 뜨는 경우 진행하지 않으셔도 됩니다.
 - <https://cofs.tistory.com/388>
 - 설치 후 'nnictl' 명령어를 Prompt에서 입력하여 에러가 발생하지 않으면 정상적으로 설치 완료입니다.
- config.yml이 세팅 된 경우 해당 파일이 있는 경로에서 nnictl create -c config.yml로 파라미터 튜닝을 수행 합니다.
 - Windows 환경 시 파이썬 환경 변수 설정이 꼬여있는 경우가 있으므로, 직접 파이썬의 경로를 지정해 주어야 합니다.
 - Anaconda Powershell Prompt에서 “get-command python” 명령어로 다음과 같이 파이썬의 경로를 찾아줍니다.
 - 파이썬의 경로를 얻은 후 config.yml 파일에 있는 command 부분에 다음과 같이 파이썬 경로를 입력해줍니다
 - Windows 환경 시 모든 파일에서 한글 주석을 제거해줍니다. 제거하지 않으면 인코딩 에러가 발생합니다.

The screenshot shows a terminal window with the following content:

```
(base) PS C:\Users\Kyle> conda activate kyle
(kyle) PS C:\Users\Kyle> get-command python
 CommandType      Name          Version      Source
-----      ----          -----      -----
 Application  python.exe    3.6.121...  C:\Users\Kyle\conda\envs\kyle\python.exe

(kyle) PS C:\Users\Kyle>
```

A red box highlights the path "C:\Users\Kyle\conda\envs\kyle\python.exe". An arrow points from this path to a callout box containing the configuration parameters for nnictl:

trial:
command: C:\Users\Kyle\conda\envs\kyle\python train.py
codeDir: .
gpuNum: 0

AutoML: NNI

설치 및 사용법

- config.yml 파일이 있는 경로에서 “nnictl create -c config.yml” 명령어로 실행하면 다음과 같은 화면이 나옵니다.

```
You can use these commands to get more information about the experiment
-----
      commands          description
1. nnictl experiment show    show the information of experiments
2. nnictl trial ls          list all of trial jobs
3. nnictl top               monitor the status of running experiments
4. nnictl log stderr        show stderr log content
5. nnictl log stdout        show stdout log content
6. nnictl stop              stop an experiment
7. nnictl trial kill        kill a trial job by id
8. nnictl --help            get help information about nnictl
-----
Command reference document https://nni.readthedocs.io/en/latest/Tutorial/Nnictl.html
```

- “nnictl stop” 명령어로 파라미터 튜닝을 멈출 수 있습니다.
튜닝을 멈추면 튜닝했던 기록이 모두 제거되니 의해주세요.

AutoML: NNI

설치 및 사용법

- 상단 Trials detail 탭에서 실험 진행 현황을 확인할 수 있습니다.
- 각 trial에 대해 Parameter, Log를 확인할 수 있는 탭이 존재 합니다.
 - 에러 발생 시 View trial error 버튼으로 어떤 에러가 발생했는지 확인할 수 있습니다.

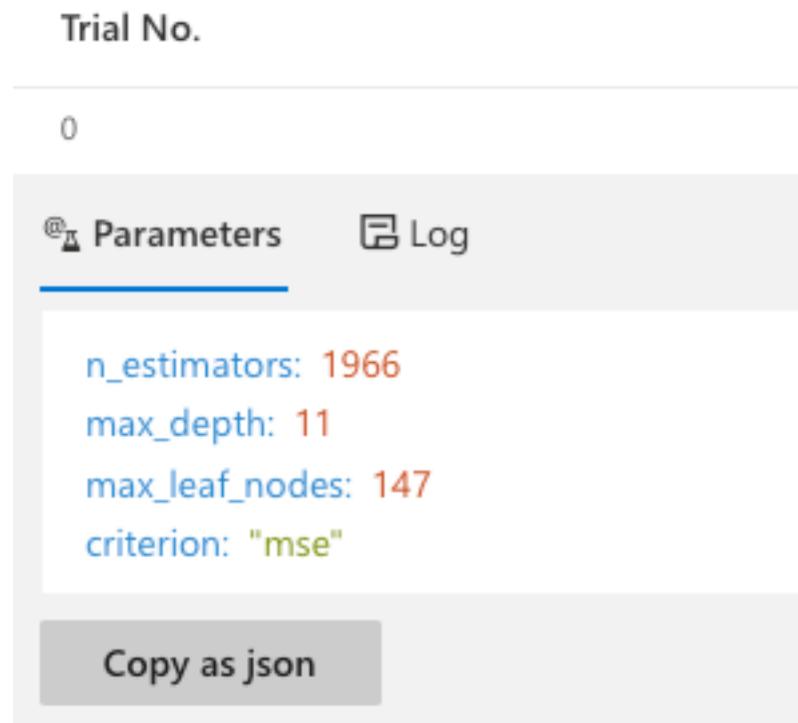
Trial No.

0

Parameters Log

n_estimators: 1966
max_depth: 11
max_leaf_nodes: 147
criterion: "mse"

Copy as json



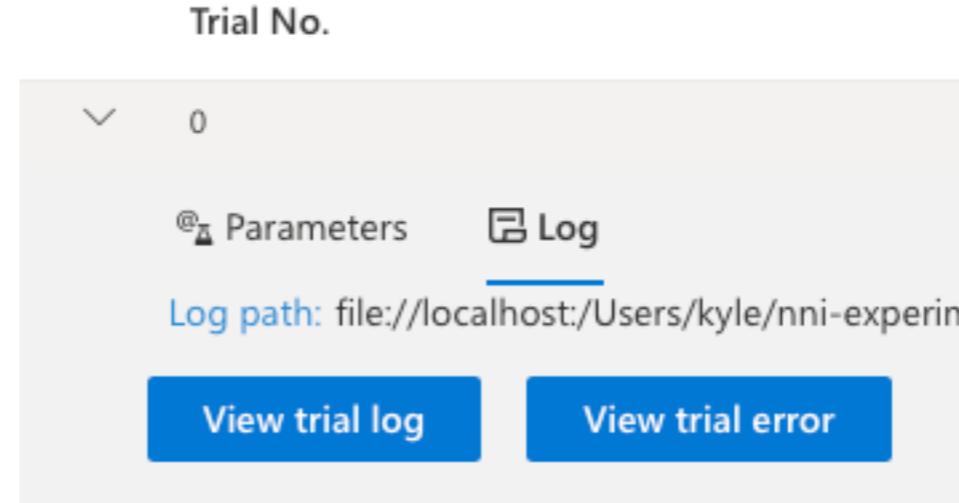
Trial No.

0

Parameters Log

Log path: file://localhost:/Users/kyle/nni-experiment/0

[View trial log](#) [View trial error](#)



실험 관리: WanDB

Weight & Biases, WanDB

- 제공하는 기능
 - 머신러닝 프로젝트 관리
 - 하이퍼 파라미터 튜닝
 - 실험 관리
 - 모델 성능 추적
 - 모델 학습 파라미터 관리

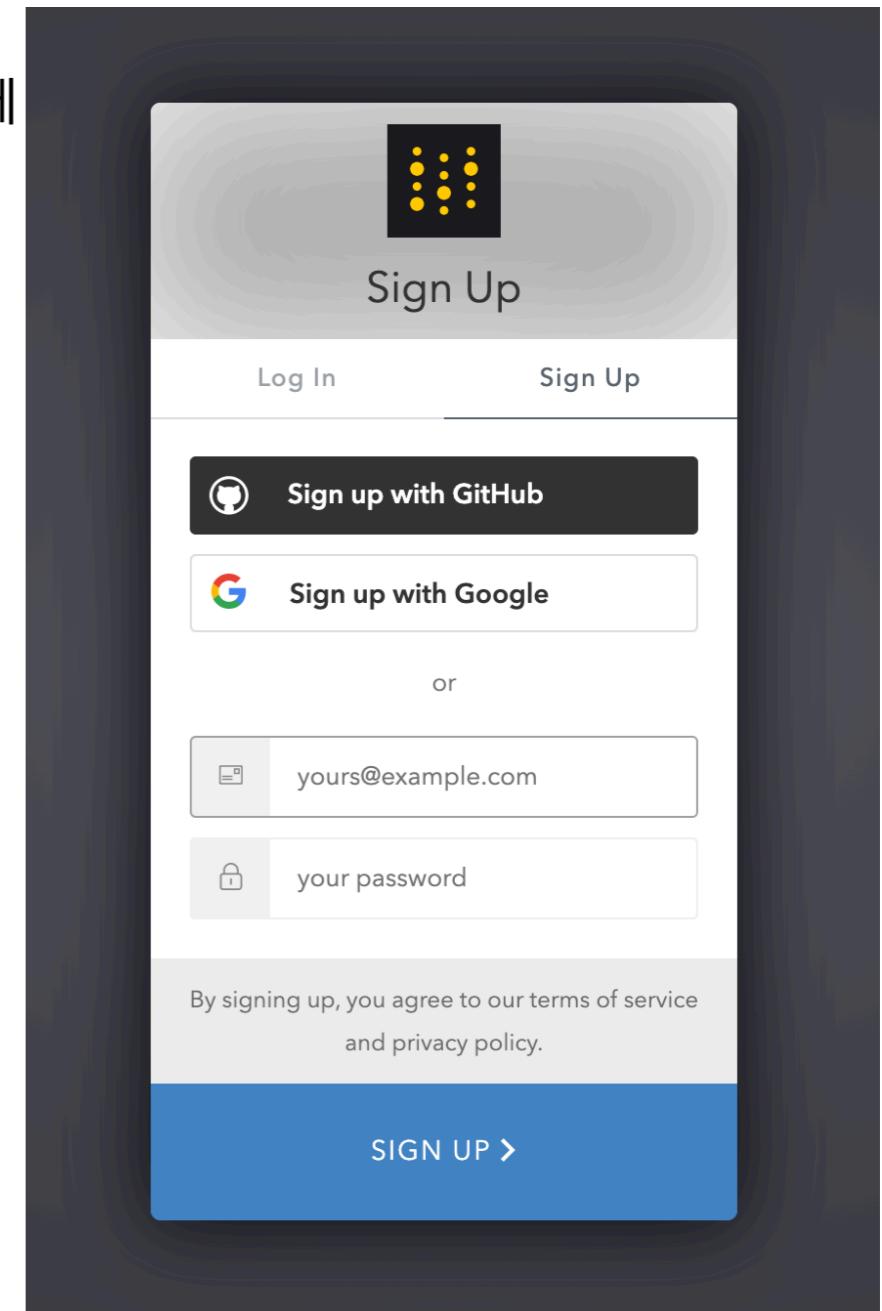
The image shows a composite view of the Weight & Biases (WanDB) platform. On the left, there is a dark-themed landing page with the text "Developer tools for machine learning" and "Build better models faster with experiment tracking, dataset versioning, and model management". It features two buttons: "SIGN UP" in a yellow box and "REQUEST DEMO" in a white box. On the right, a light-themed dashboard is displayed, showing the results of a "Hyperparameter Sweep". The dashboard includes:

- A main plot titled "Results of Hyperparameter Sweep 5" showing the relationship between hyperparameters (n_params, activation, batch_size, shape) and validation accuracy (val-full/accuracy). The accuracy ranges from 0.90 to 0.99.
- A "Parameter importance with respect to val-full" section with a bar chart showing the importance of various parameters like n_params, budget, parameter_budget, Runtime, lr, batch_size, activation.value_relu, activation.value_tanh, activation.value_sig..., and shape.value_deep.
- A "loss/val-batch" line chart showing training progress over 300 epochs.
- A summary section titled "Best Model by Val Acc" showing a large value of 0.981.

실험 관리: WanDB

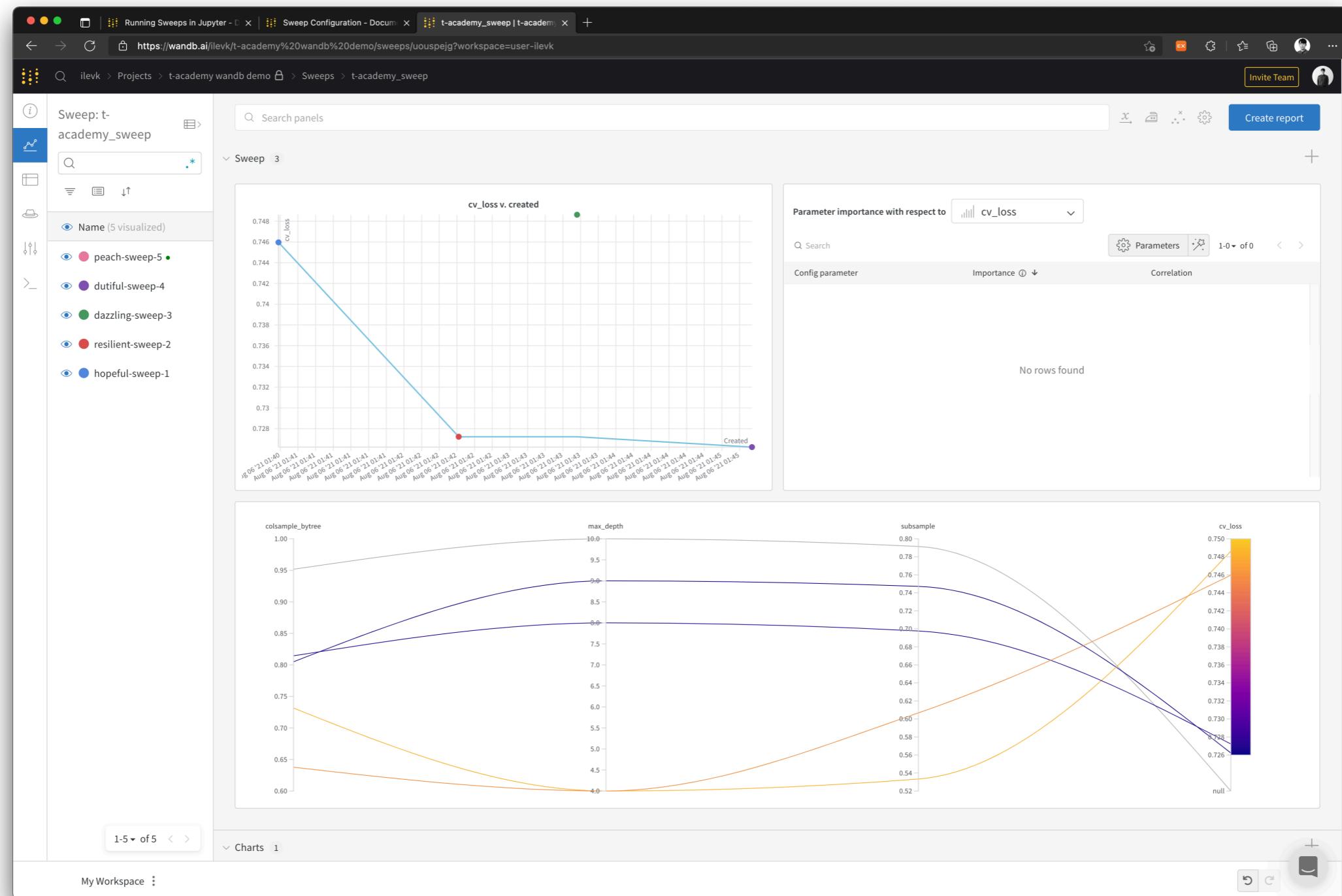
설치 및 사용법

1. wandb.ai 사이트 회원 가입
 - Github 또는 Google 계정을 이용해 가입할 수 있습니다.
 - wandb 사이트에서 프로젝트 관리 및 실험 결과를 확인할 수 있기 때문에 회원가입은 반드시 해야합니다.
2. 'pip install wandb' 커맨드로 wandb 라이브러리 설치
3. 'wandb login' 커맨드로 wandb 서버와 연동을 진행합니다.
 - 초기 한번만 세팅, 로컬에 API Key 정보를 저장합니다.



실험 관리: WanDB

파라미터 튜닝 Dashboard



Q&A