Hydra

December 12, 2022

Hydra

In this document, we are going to explain Hydra, open source python framework for hiearchical configuration.

Contents

1. What is Hydra?

Official Website

Hydra is an open-source Python framework that simplifies the development of research and other complex applications. The key feature is the ability to dynamically create a hierarchical configuration by composition and override it through config files and the command line. The name Hydra comes from its ability to run multiple similar jobs - much like a Hydra with multiple heads.

Summary

- ELI5(Explain Like I'm 5) excerpted from Youtube comment : ArgParse on steriods, with multiple functionalities such as config files,changing ,saving config files and much more .

2. Features

- Hierarchical configuration composable from multiple sources
  - using .yaml file
- Configuration can be specified or overridden from the command line
  - also dynamically change variables using OmegaConf's variable interpolation syntax

```yaml
# config.yaml
defaults:
  - dataset: imagenet
  - model: alexnet
  - optional dataset_model: ${dataset}_${model}
```

- Dynamic command line tab completion
- Run your application locally or launch it to run remotely
  - using Ray Launcher plugin, run the code on AWS
- Run multiple jobs with different arguments with a single command
  - Run multiple jobs in parallel mode

3. Usuage

This section might include personal opinions.

- Configurations
  - Hydra is a powerful configuration framework. Not only for deep learning, but also it can be applicable to other applications such as Databases, Web Applications.
- Management
  - provide easy logging
  - provide detail configuration settings by providing
    - config.yaml: A dump of the user specified configuration
    - hydra.yaml: A dump of the Hydra configuration
    - overrides.yaml: The command line overrides used
- Multi-run
  - Useful for parameter searching

4. Conclusion

This section includes personal opinions about using Hydra

- When using Huggingface Transformers, it provides HfArgumentParser for configurations such as
  - ModelArguments, DataTrainingArguments, TrainingArguments
- HfArgumentParser is simliar to Hydra, specified for Transformers library.
  - Of course, one can make hydra version of HfArgumentParser and HfArgumentParser version of hydra.
  - HfArgumentParser, Hydra all based on using dataclasses in python
  - It depends on user's preference
- But for the other cases such as setting-up the applications, databases or using facebook library
  - Hydra looks like a better option
- Recommendation
  - For huggingface transformers, use HfArgumentParser
  - For other cases, use Hydra
  - Or user's choice

Reference

Hydra
Explain Like I'm 5: Hydra
Configuration Management For Data Science Made Easy With Hydra
Github