# 第六次课笔记（作业截图附后）

2024年1月28日  13:55

环境准备：

## 🔗 安装

### 面向GPU的环境安装

```
conda create --name opencompass --clone=/root/share/conda_envs/internlm-base
source activate opencompass
git clone https://github.com/open-compass/opencompass
cd opencompass
pip install -e .
```

有部分第三方功能,如代码能力基准测试 Humaneval 以及 Llama格式的模型评测,可能需要额外步骤才能正常运行，如需评测，详细步骤请参考安装指南。

### 数据准备

```
# 解压评测数据集到 data/ 处
cp /share/temp/datasets/OpenCompassData-core-20231110.zip /root/opencompass/
unzip OpenCompassData-core-20231110.zip

# 将会在opencompass下看到data文件夹
```

### 查看支持的数据集和模型

```
# 列出所有跟 internlm 及 ceval 相关的配置
python tools/list_configs.py internlm ceval
```

显示结果：

```
(opencompass) root@intern-studio:~/opencompass# python tools/list_configs.py internlm ceval
+----------------------------+--------------------------------------------------------------+
| Model                      | Config Path                                                  |
+----------------------------+--------------------------------------------------------------+
| hf_internlm2_20b           | configs/models/hf_internlm/hf_internlm2_20b.py               |
| hf_internlm2_7b            | configs/models/hf_internlm/hf_internlm2_7b.py                |
| hf_internlm2_chat_20b      | configs/models/hf_internlm/hf_internlm2_chat_20b.py          |
| hf_internlm2_chat_7b       | configs/models/hf_internlm/hf_internlm2_chat_7b.py           |
| hf_internlm_20b            | configs/models/hf_internlm/hf_internlm_20b.py                |
| hf_internlm_7b             | configs/models/hf_internlm/hf_internlm_7b.py                 |
| hf_internlm_chat_20b       | configs/models/hf_internlm/hf_internlm_chat_20b.py           |
| hf_internlm_chat_7b        | configs/models/hf_internlm/hf_internlm_chat_7b.py            |
| hf_internlm_chat_7b_8k     | configs/models/hf_internlm/hf_internlm_chat_7b_8k.py         |
| hf_internlm_chat_7b_v1_1   | configs/models/hf_internlm/hf_internlm_chat_7b_v1_1.py       |
| internlm_7b                | configs/models/internlm/internlm_7b.py                       |
| ms_internlm_chat_7b_8k     | configs/models/ms_internlm/ms_internlm_chat_7b_8k.py         |
+----------------------------+--------------------------------------------------------------+

+----------------------------+--------------------------------------------------------------+
| Dataset                    | Config Path                                                  |
+----------------------------+--------------------------------------------------------------+
| ceval_clean_ppl            | configs/datasets/ceval/ceval_clean_ppl.py                   |
| ceval_gen                  | configs/datasets/ceval/ceval_gen.py                         |
| ceval_gen_2daf24           | configs/datasets/ceval/ceval_gen_2daf24.py                  |
| ceval_gen_5f30c7           | configs/datasets/ceval/ceval_gen_5f30c7.py                  |
| ceval_ppl                  | configs/datasets/ceval/ceval_ppl.py                         |
| ceval_ppl_578f8d           | configs/datasets/ceval/ceval_ppl_578f8d.py                  |
| ceval_ppl_93e5ce           | configs/datasets/ceval/ceval_ppl_93e5ce.py                  |
| ceval_zero_shot_gen_bd40ef | configs/datasets/ceval/ceval_zero_shot_gen_bd40ef.py        |
+----------------------------+--------------------------------------------------------------+
```

启动评测：

```
(opencompass) root@intern-studio:~/opencompass# pyth
on run.py --datasets ceval_gen --hf-path /share/temp
/model_repos/internlm-chat-7b/ --tokenizer-path /sha
re/temp/model_repos/internlm-chat-7b/ --tokenizer-kw
args padding_side='left' truncation='left' trust_rem
ote_code=True --model-kwargs trust_remote_code=True
device_map='auto' --max-seq-len 2048 --max-out-len 1
6 --batch-size 4 --num-gpus 1 --debug
```

评测结果：

```
dataset                                         version    metric          mode     opencompass.models.huggingface.HuggingFace_model_repos_internlm-chat-7b
--------------------------------------------    -------    -------------   ------   -----------------------------------------------------------------------
ceval-computer_network                          db9ce2     accuracy        gen                                                                        31.58
ceval-operating_system                          1c2571     accuracy        gen                                                                        36.84
ceval-computer_architecture                     a74dad     accuracy        gen                                                                        28.57
ceval-college_programming                       4ca32a     accuracy        gen                                                                        32.43
ceval-college_physics                           963fa8     accuracy        gen                                                                        26.32
ceval-college_chemistry                         e78857     accuracy        gen                                                                        16.67
ceval-advanced_mathematics                      ce03e2     accuracy        gen                                                                        21.05
ceval-probability_and_statistics                65e812     accuracy        gen                                                                        38.89
ceval-discrete_mathematics                      e894ae     accuracy        gen                                                                        18.75
ceval-electrical_engineer                       ae42b9     accuracy        gen                                                                        35.14
ceval-metrology_engineer                        ee34ea     accuracy        gen                                                                        50
ceval-high_school_mathematics                   1dc5bf     accuracy        gen                                                                        22.22
ceval-high_school_physics                       adf25f     accuracy        gen                                                                        31.58
ceval-high_school_chemistry                     2ed27f     accuracy        gen                                                                        15.79
ceval-high_school_biology                       8e2b9a     accuracy        gen                                                                        36.84
ceval-middle_school_mathematics                 bee8d5     accuracy        gen                                                                        26.32
ceval-middle_school_biology                     86817c     accuracy        gen                                                                        61.9
ceval-middle_school_physics                     8accf6     accuracy        gen                                                                        63.16
ceval-middle_school_chemistry                   167a15     accuracy        gen                                                                        60
ceval-veterinary_medicine                       b4e08d     accuracy        gen                                                                        47.83
ceval-college_economics                         f3f4e6     accuracy        gen                                                                        41.82
ceval-business_administration                   c1614e     accuracy        gen                                                                        33.33
ceval-marxism                                   cf874c     accuracy        gen                                                                        68.42

ceval-high_school_politics                      5c0de2     accuracy        gen                                                                        26.32
ceval-high_school_geography                     865461     accuracy        gen                                                                        47.37
ceval-middle_school_politics                    5be3e7     accuracy        gen                                                                        52.38
ceval-middle_school_geography                   8a63be     accuracy        gen                                                                        58.33
ceval-modern_chinese_history                    fc01af     accuracy        gen                                                                        73.91
ceval-ideological_and_moral_cultivation         a2aa4a     accuracy        gen                                                                        63.16
ceval-logic                                     f5b022     accuracy        gen                                                                        31.82
ceval-law                                       a110a1     accuracy        gen                                                                        25
ceval-chinese_language_and_literature           0f8b68     accuracy        gen                                                                        30.43
ceval-art_studies                               2a1300     accuracy        gen                                                                        60.61
ceval-professional_tour_guide                   4e673e     accuracy        gen                                                                        62.07
ceval-legal_professional                        ce8787     accuracy        gen                                                                        39.13
ceval-high_school_chinese                       315705     accuracy        gen                                                                        63.16
ceval-high_school_history                       7eb30a     accuracy        gen                                                                        70
ceval-middle_school_history                     48ab4a     accuracy        gen                                                                        59.09
ceval-civil_servant                             87d061     accuracy        gen                                                                        53.19
ceval-sports_science                            70f27b     accuracy        gen                                                                        52.63
ceval-plant_protection                          8941f9     accuracy        gen                                                                        59.09
ceval-basic_medicine                            c409d6     accuracy        gen                                                                        47.37
ceval-clinical_medicine                         49e82d     accuracy        gen                                                                        40.91
ceval-urban_and_rural_planner                   95b885     accuracy        gen                                                                        45.65
ceval-accountant                                002837     accuracy        gen                                                                        26.53
ceval-fire_engineer                             bc23f5     accuracy        gen                                                                        22.58
ceval-environmental_impact_assessment_engineer  c64e2d     accuracy        gen                                                                        64.52
ceval-tax_accountant                            3a5e3c     accuracy        gen                                                                        34.69
ceval-physician                                 6e277d     accuracy        gen                                                                        40.82
ceval-stem                                      -          naive_average   gen                                                                        35.09
ceval-social-science                            -          naive_average   gen                                                                        52.79
ceval-humanities                                -          naive_average   gen                                                                        52.58
ceval-other                                     -          naive_average   gen                                                                        44.36
ceval-hard                                      -          naive_average   gen                                                                        23.91
ceval                                           -          naive_average   gen                                                                        44.16
01/28 14:25:39 - OpenCompass - INFO - write summary to /root/opencompass/outputs/default/20240128_135731/summary/summary_20240128_135731.txt
01/28 14:25:39 - OpenCompass - INFO - write csv to /root/opencompass/outputs/default/20240128_135731/summary/summary_20240128_135731.csv
(opencompass) root@intern-studio:~/opencompass#
```