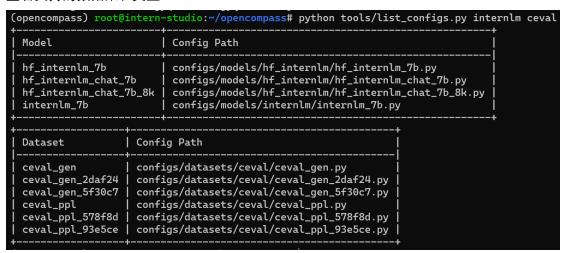
基础作业:

使用 OpenCompass 评测 InternLM2-Chat-7B 模型在 C-Eval 数据集上的性能 Github 连不上,用了 Gtee。

查看支持的数据集和类型:



过程中: *注意把 batchsize 改小

```
Loading checkpoint shards: 100% | 01/22 20:40:47 - OpenCompass - INFO - Start inferencing [opencompass.models.huggingface.Huggingface_model_repos_internIm -chat-Tb/ceval-college_economics] | 55/55 [00:00<00:00, 1325785.75it/s] | 100% | 14/14 [00:35<00:00, 1325785.75it/s] | 100% | 14/14 [00:35<00:00, 2.54s/it] | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12 | 16/12
```

评测结果:

ceval-urban_and_rural_planner	45.65	95b885	accuracy	gen
ceval-accountant		002837	accuracy	gen
ceval-fire_engineer	26.53	bc23f5	accuracy	gen
ceval-environmental_impact_as		c64e2d	accuracy	gen
ceval-tax_accountant	64.52	3a5e3c	accuracy	gen
ceval-physician	34.69	6e277d	accuracy	gen
ceval-stem	44.9		naive_average	gen
ceval-social-science	35.87		naive average	
	52.2		- ,	
ceval-humanities	52.1		naive_average	
ceval-other	44.73		naive_average	gen
ceval-hard	24.57		naive_average	gen
ceval	44.32		naive_average	gen
01/22 23:51:40 - OpenCompass - <u>INFO</u> - write summary to /root/opencompass/outputs/default/20240122_224139/summary _20240122_224139.txt				
01/22 23:51:40 - OpenCompass - <u>INFO</u> - write csv to /root/opencompass/outputs/default/20240122_224139/summary/summary_202 40122 224139.csv				
(opencompass) root@intern-studio:~/opencompass#				
Mccompassl 0:bash*			· · · · · · · · · · · · · · · · · · ·	"intern-studio" 10:51 73-Jan-74