

OpenSeek 大模型挑战赛--初赛技术报告

ccabcca06

一、任务背景

初赛阶段以 Openseek-small-v1（一个总参数量 1.4B，激活参数量 0.4B 的混合专家模型-MoE）为起点，基于主办方提供的 OpenSeek-Pretrain-100B 多源异构数据集，依托 FlagScale 训练框架，开展预训练实验。训练完成后，使用 lighteval 评估框架，在涵盖常识推理、世界知识、数学计算及中英文理解能力的 12 个权威评测基准（Hellaswag, Winogrande, PIQA, SIQA, OpenBookQA, ARC-easy, ARC-challenge, CommonsenseQA, MMLU-ProMC, BoolQ, CMMLU, C-Eval）上对模型性能进行了全面评估。

OpenSeek-Pretrain-100B 是一个大规模、高质量、多样化的开源预训练语料库，总规模达到 1000 亿 tokens（约 100B Tokens）。100B 数据集主要包括以下几类子集：

1. 通用网络文本：主要包括 Nemotron-CC 和 zh_cc。Nemotron-CC 是 NVIDIA Nemotron 数据集的 Common Crawl 部分，zh_cc 是从 Common Crawl 中提取的中文网页数据；

2. 合成思维链数据（CoT Synthesis）：包含了由教师模型（如 GPT-4）生成的大量推理过程数据。这些数据被分为“第一代”（cot_synthesis）和“改进的第二代”（cot_synthesis2），第二代

通常质量更高、更复杂；

3.领域专用数据（学术、代码、书籍等）：如来自知名的学术预印本网站 [arXiv.org](https://arxiv.org) 的大量科学、技术、工程和数学（STEM）领域学术论文（通常是 **LaTeX** 源码或提取的文本）内容；来自多个开源代码库的数据，例如 **GitHub** 上的公共项目，经过过滤的高质量代码片段和完整文件；各种电子书项目（如 **Project Gutenberg**）和开源书籍库；来自数学教材、学术论文（如 **Arxiv** 数学板块）、数学竞赛题目（如 **AMC, AOPS**）、包含约 4000 万篇开放获取学术论文的 **Pes2o** 学术文本预训练数据集以及 **Stack Exchange** 网络下的问答平台（如 **Stack Overflow, Mathematics, Physics** 等）的高质量的编程、技术和学术问答对。

二、整体思路

首先，我们对 **OpenSeek-Pretrain-100B** 的各类子集进行了逐一分析，根据其内容形式、知识密度和预估难度将其划分为四个类别：
基础语言与通用知识（Foundation）：此类数据构成模型的语言理解和世界知识基础，包括：

Nemotron-CC, zh_cc, wiki：高质量的通用网页文本和中英文百科，语言规范，知识覆盖面广，难度适中。

Books：大量书籍数据，包含丰富的叙事性、论述性长文本，有助于模型建模长程依赖和复杂语言结构。

Pes2o, arxiv, Stack：学术论文、技术问答与社区讨论数据，包含大

量解释性文本，逻辑性较强。

代码数据 (Code)： 此类数据高度结构化，逻辑严谨。Code 子集用于培养模型的逻辑思维、精确性和结构化生成能力。

数学数据 (Math)： Math 子集包含数学符号、公式和推理过程，是最高难度的数据之一，要求模型具备极强的抽象和符号推理能力。

思维链数据 (CoT Synthesis)： 此类数据（包括 `cot_synthesis*` 系列）并非原始数据，而是由更高级模型生成的、包含推理步骤 (Chain-of-Thought) 的合成数据。这些数据显式地展示了如何一步步思考和解决问题，为模型提供了元认知的范本。

课程学习 (Curriculum learning, CL) 是一种训练策略，模仿人类的学习过程，主张让模型先从容易的样本开始学习，并逐渐进阶到复杂的样本和知识。CL 策略在计算机视觉和自然语言处理等多种场景下，在提高各种模型的泛化能力和收敛率方面表现出了强大的能力。

基于 OpenSeek-Pretrain-100B 的各个子集的特性，我们采用从易到难的课程学习训练设计，将预训练过程划分为三个阶段：

第一阶段：基础语言建模 (Foundation Modeling)

目标： 让模型掌握流畅、规范的语言生成能力，并建立起基本的世界知识图谱。

数据配比： 重点倾斜至基础语言与通用知识类别（如 Nemotron-CC, Wiki, Books, zh_cc）。此阶段占比最高（50%以上），确保模型打下坚实的语言基础。

设计思路： 避免过早接触高难度或特殊领域数据（如代码、数学）导

致模型混淆或产生“知识幻觉”。纯净、规范的语言环境有助于模型形成正确的语法、句法和语义概念。

第二阶段：逻辑与推理能力专项提升 (Specialization)

目标：在打好语言基础后，引入代码和数学数据，专项提升模型的逻辑推理、结构化思维和精确计算能力。

数据配比：显著提升代码和数学数据的采样比例（合计提升至~30%）。同时，开始引入思维链数据（~10%），让模型学习如何拆解复杂问题。

设计思路：此时模型已具备良好的语言理解能力，能够更好地理解代码注释、数学问题描述。思维链数据作为“脚手架”，引导模型模仿推理过程，降低直接学习代码和数学的难度。

第三阶段：综合推理与知识融汇 (Integration)

目标：促进不同领域知识的融合，激发模型的涌现能力，使其能灵活运用所学知识解决复杂、开放的问题。

数据配比：最大化思维链数据的比例（提升至~20%+）。其他数据的比例相对降低但保持均衡，确保知识的广度和多样性。

设计思路：高比例的思维链数据相当于让模型持续进行“高强度的推理演练”，迫使它将前两个阶段学到的语言、知识和逻辑技能结合起来，模仿并内化复杂的推理模式。这是模型性能实现飞跃的关键阶段。

三、实验记录

3.1 基础模型 baseline

使用 Openseek-small-v1 模型，按照文档要求进行 tokenizer 替换处理后，直接进行评估任务，最终得分为 36.57,具体如下：



Openseek-small-v1 基线分数

3.2 训练 pipeline

基于 FlagScale 镜像默认配置，直接运行默认训练脚本，训练完成后得分为 30.75，具体如下：

运行反馈															✕	
% Total	% Received	% Xferd	Average	Speed	Time	Time	Time	Current								
			Dload	Upload	Total	Spent	Left	Speed								
0	0	0	0	0	0	0	0	0	0100	272k	100	272k	0	0	639k	0
640k																
HellaSwag 35.27																
ARC(Average) 36.15																
PIQA 62.95																
MMLU(cloze) 27.86																
CommonsenseQA 22.6																
TriviaQA 3.13																
Winograde 49.96																
OpenBookQA 30.8																
GSM8K(5-shot) 0.99																
SIQA 41.71																
CEval 28.83																
CMMLU 28.8																

FlagScale 默认训练配置得分

3.3 课程学习

基于原始网友爬虫数据(Nemotron-CC-high-actual-actual-high, Nemotron-CC-high-actual-actual-mid), 进行第一阶段课程学习训练后，最终评测得分为 34.48，具体如下：

运行反馈															✕	
% Total	% Received	% Xferd	Average	Speed	Time	Time	Time	Current								
			Dload	Upload	Total	Spent	Left	Speed								
0	0	0	0	0	0	0	0	0	0	1	272k	1	3510	0	0	13788
0:00:20 13764100 272k 100 272k 0 0 640k 0 0:00:20 640k																
HellaSwag 45.32																
ARC(Average) 43.73																
PIQA 70.35																
MMLU(cloze) 31.8																
CommonsenseQA 32.35																
TriviaQA 10.51																
Winograde 50.28																
OpenBookQA 31.8																
GSM8K(5-shot) 1.59																
SIQA 42.68																
CEval 26.7																
CMMLU 26.61																

一阶段课程学习训练得分

3.4 结论

由于时间及算力资源限制原因，未能完成全部的课程学习训练。从初步实验可以看出，对比全量数据混合，课程学习设置确实有比较明显的提升效果。但是均低于 **benchmark** 的结果显示课程设置和数据选择及配比还有更大的提升空间。