

Team011の取り組み

日立製作所 研究開発グループ 友成光 竹下孔喜 角掛正弥 森下皓文

HITACHI

概要



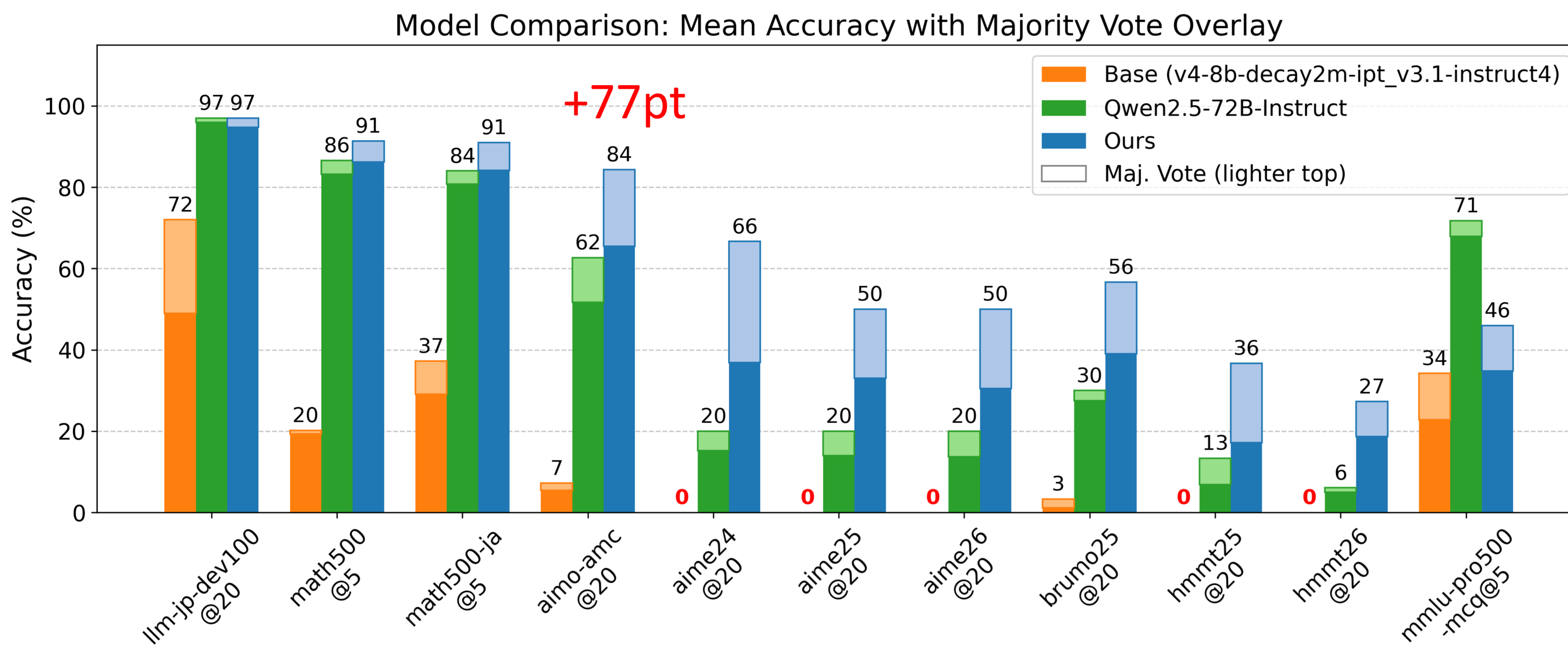
特にコンテキスト拡張に注力

実験設定

項目	継続事前学習 (CPT)	教師ありファインチューニング (SFT)
主な目的	知識・表現の獲得	推論能力の強化
データ	nvidia/Nemotron-CC-Math-v1 の 4plus サブセット	OpenDataArena/ODA-Math-460k (回答はAM-Thinking-v1で合成)
前処理	品質フィルタリングとパッキング	トランケーション
トークン数	約20B	約3B
言語	英語	問題・回答ともに英語
学習元モデル	v4-8b-decay2m-ipt_v3.1-instruct4 (8B)	CPTモデル
コンテキスト長	4k → 16k YaRN [Peng et al., 2023] を用いて拡張	16k
オプティマイザと学習率	AdamW 1e-4	AdamW 5e-5
学習率スケジュール	warmup 5% → cosine decay (5e-6まで)	warmup 5%
エポック数	1	2
バッチサイズ	64	64

- データセットの選定：複数の問題データをもとに回答合成モデルを変えながら小規模学習し、ベンチマークのスコアを比較して決定
- ハイパーパラメータの選定：olmo3などのテクニカルペーパーと小規模実験をもとに決定

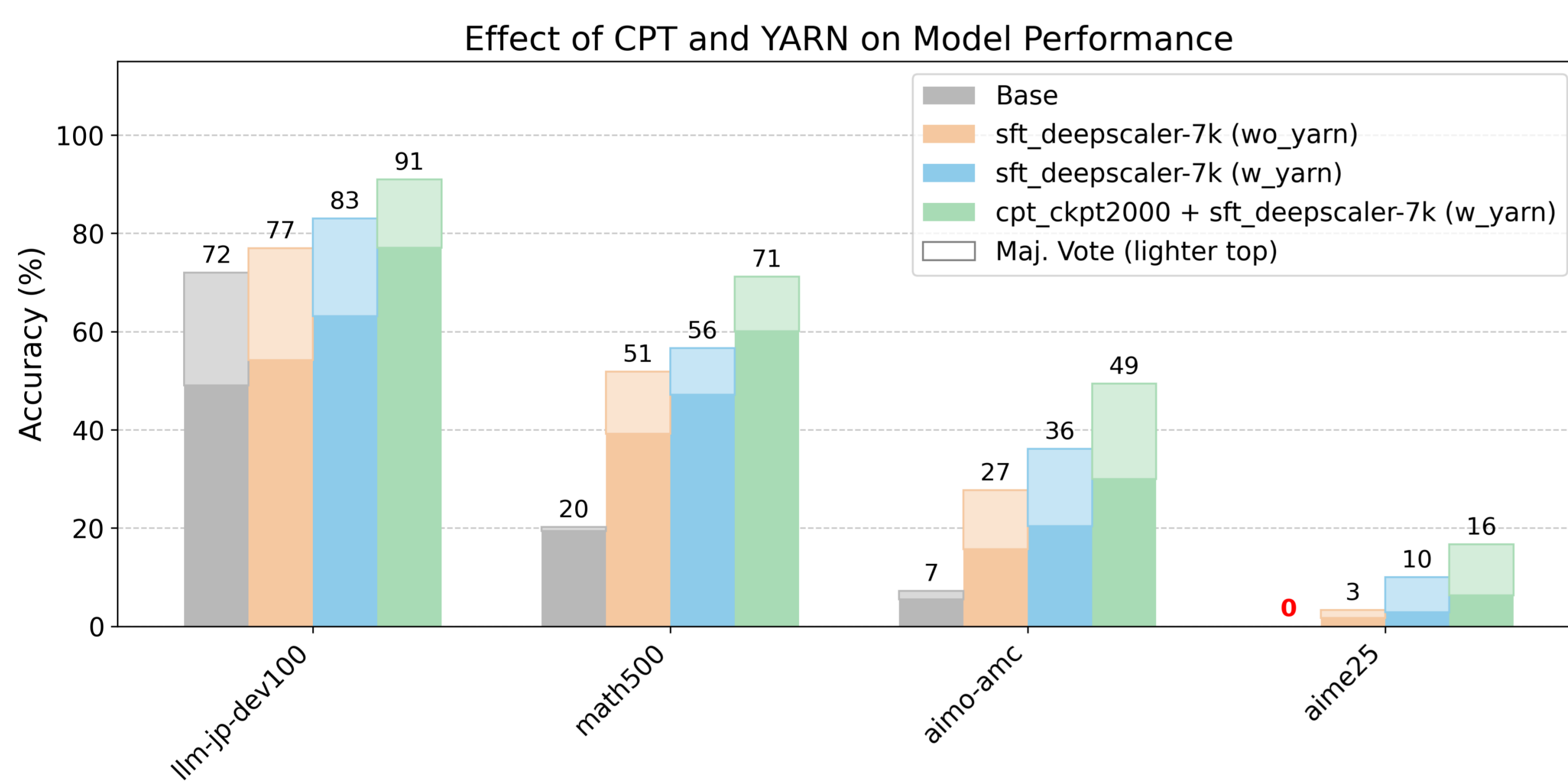
評価結果



多様なベンチマークでスコア向上

- 開発データで**25pt**
- 英語で学習したが、日本語のMATH500スコアが英語と同水準
➡ 英語で獲得する能力は事前学習でよく見た言語（日本語）に転移しやすい可能性 [Barua et al., 2025]
- 難関ベンチマークでは、最大**77pt**
- 多様なドメインの問題を含むMMLU-Proで**12pt**

分析



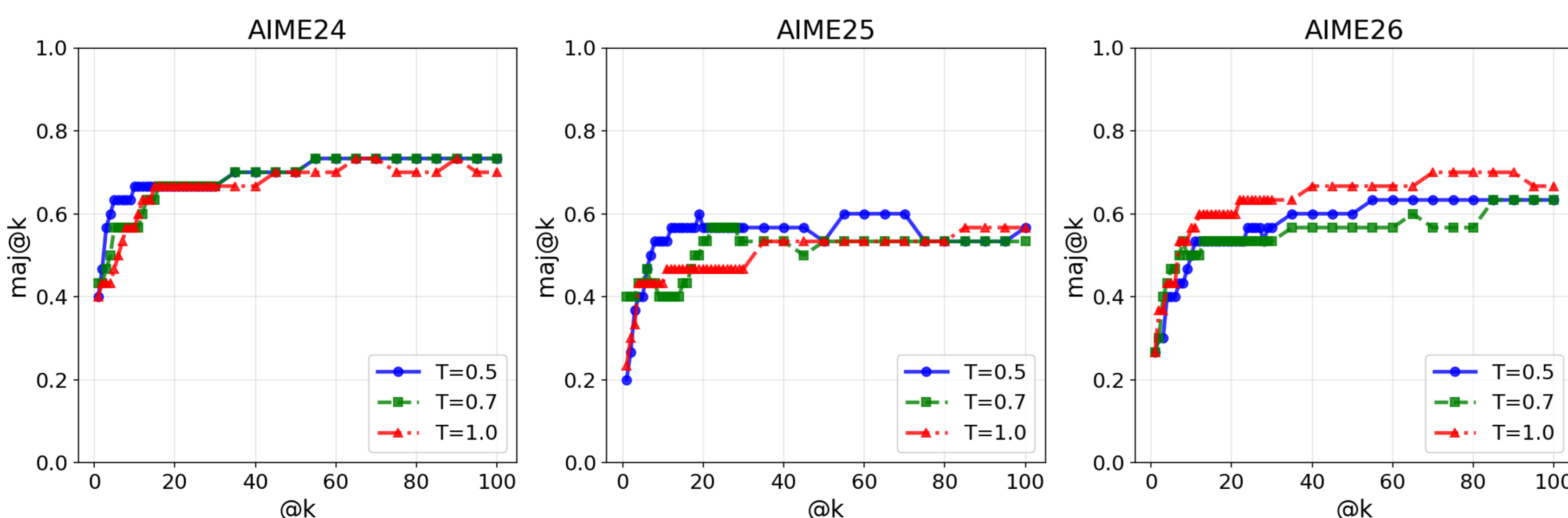
YaRNと継続事前学習の効果を調べる小規模実験

- ロングコンテキストのSFTデータとして、DeepScaleRのサブセットの回答をQwen3-30b-a3bで合成
- 少量データの学習で以下を比較
 - 直接外挿によるSFT
 - YaRNを用いたSFT
 - YaRNを用いたCPT+SFT

➡ 一貫してYaRNとCPTによりスコア向上

その他の知見と工夫点

- 継続事前学習のデータフィルタリング
➡ 質の高さが重要、短文も混ぜるべき
- Tool Integrated Reasoning [Gou et al., 2023] (コードを実行しながら推論)
➡ 開発データでは効果が薄い
- 誤りパターン分析
➡ どういうミスが多いかをLLMで分析・改善策提案により分析効率化
- 強化学習
➡ SFT (蒸留) に比べるとコストが高い。が汎化しやすい可能性 [Huan et al., 2025]



- 提出モデルで多数決の数を増やすと約@30まで順調にスコア向上
- 温度パラメータを変えたときの多数決スコアに一貫した傾向はみられない

参考文献 Huan et al., "Does Math Reasoning Improve General LLM Capabilities? Understanding Transferability of LLM Reasoning", 2025.

Gou et al., ToRA: A Tool-Integrated Reasoning Agent for Mathematical Problem Solving, arXiv:2309.17452, 2023.

Barua et al., Long Chain-of-Thought Reasoning Across Languages, arXiv:2508.14828, 2025.

Copyright © Hitachi, Ltd. 2025. All rights reserved.