

自己生成データを活用した事後学習による大規模言語モデルの性能検証

吉田 希世, 齋藤 由佳, 津田 純花, 内藤 悠, 朱 灝丞, 吉田 倖 (東北大),
赤間 怜奈 (東北大/国語研/理研), 鈴木 潤 (東北大/理研/NII LLMC)

yoshida.kise.t1@dc.tohoku.ac.jp

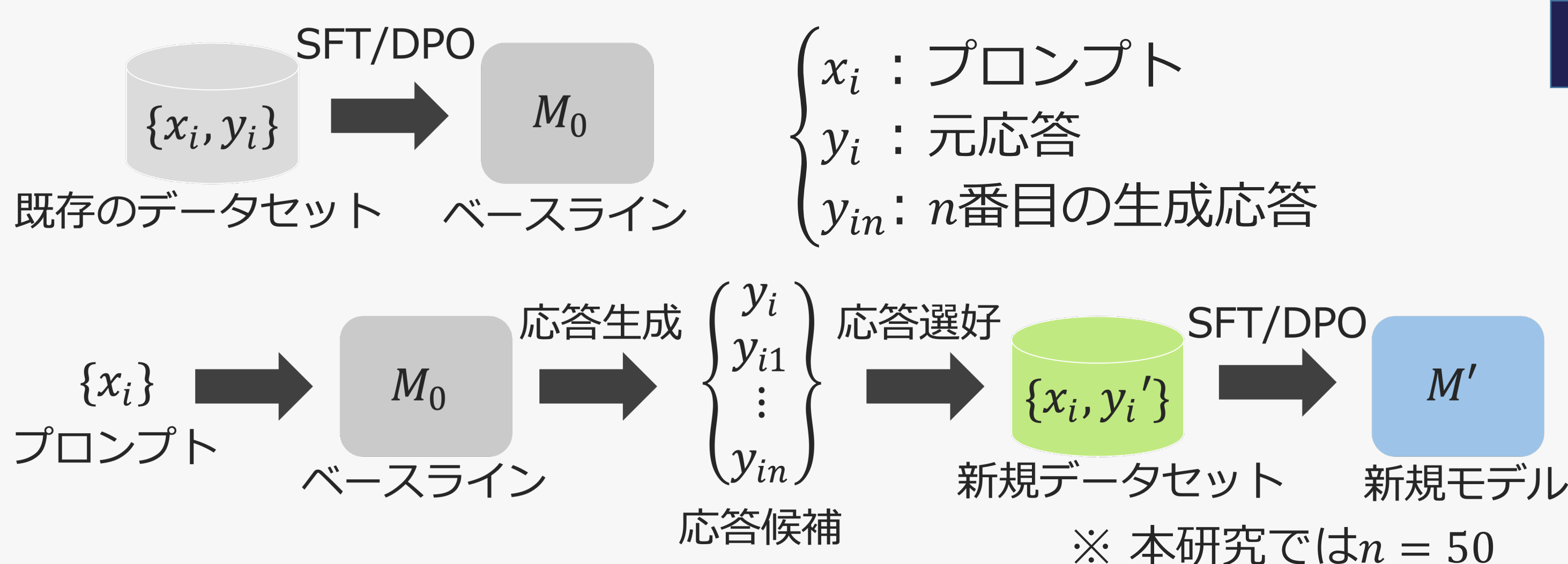
S5-P15

背景／動機

- 先行研究では、学習対象とは別の言語モデルが生成したデータを用いた事後学習によって性能向上
ex.) Alpaca 7B [Taori+ '23]
- 性能のいいモデルの能力を移しているだけでそのモデルに性能が依存してしまうのではないか
- 学習対象のモデル自身のみで性能向上を図りたい

学習手順

- 既存のデータセットからプロンプト x_i を抽出
- ベースライン M_0 に x_i を与えて応答を n 個生成
- n 個の応答から新たな応答を選好
- 新規データセットを用いて M_0 を学習



実験設定

ベースライン:

- SFT/DPO済みのllm-jp-3 (150m/980m/1.8b/3.7b)

データセット:

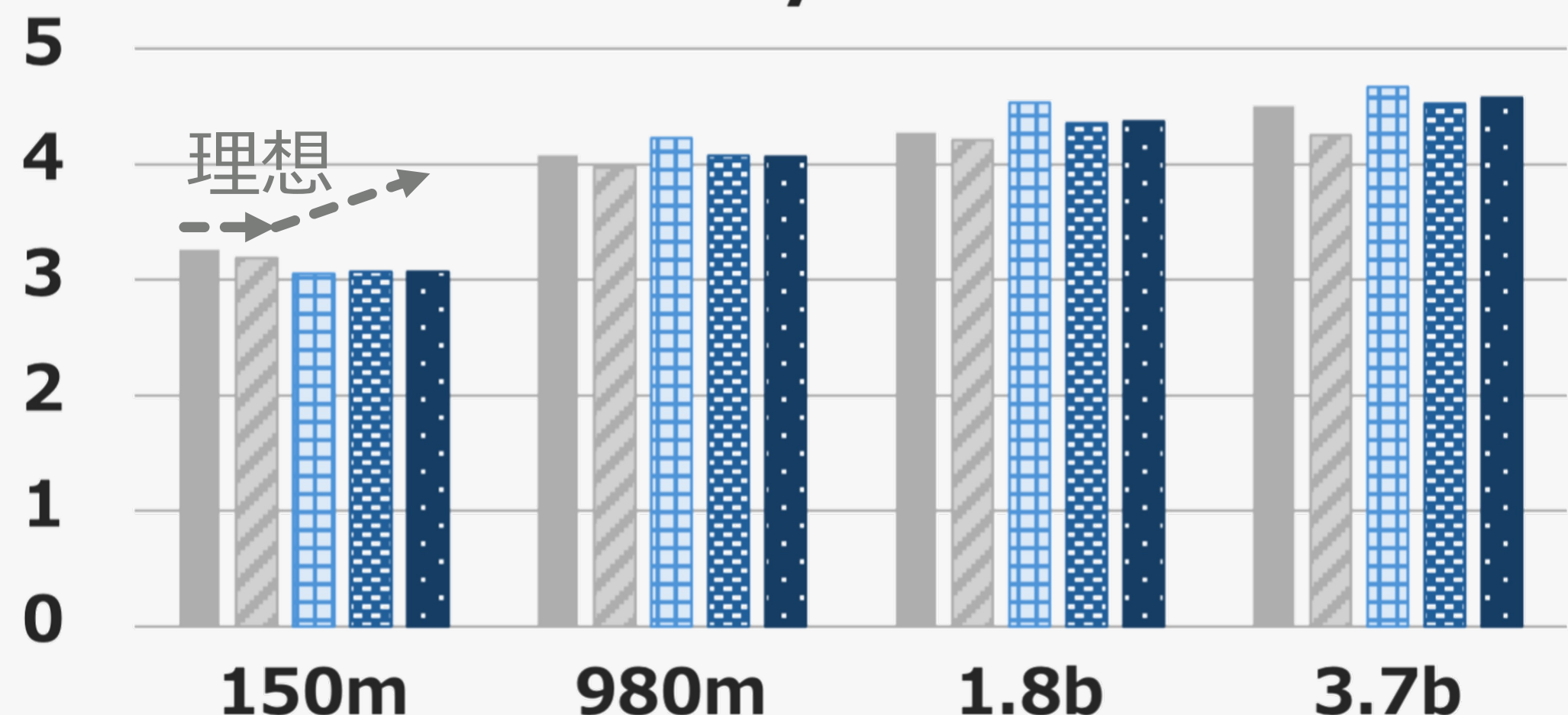
- llm-jp-3-instruct3 [勝又+ '25] の学習データ

実験結果／分析

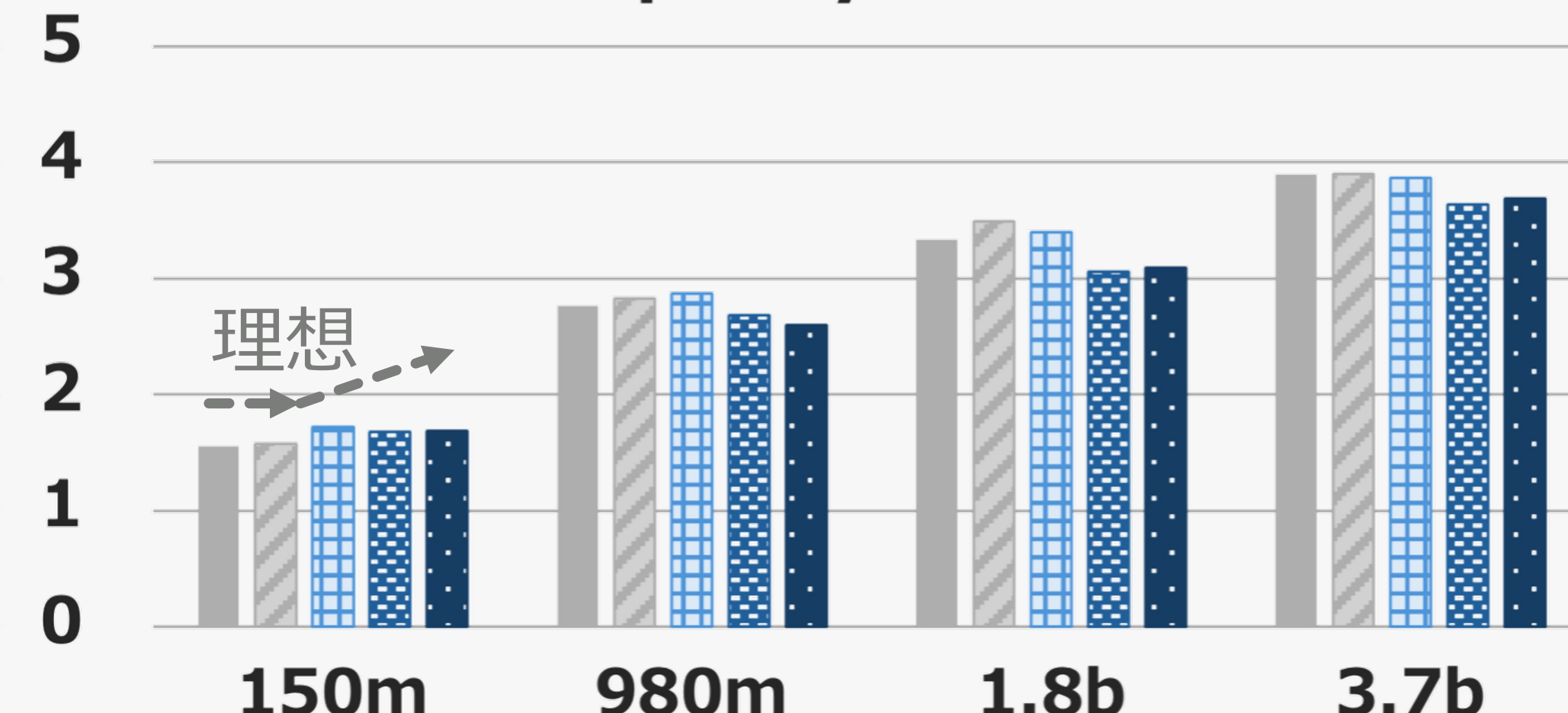
■ ベースライン ■ 手法① ■ 手法② ■ 手法③ ■ 手法④

llm-jp-judgeでの評価結果

safety: 安全性

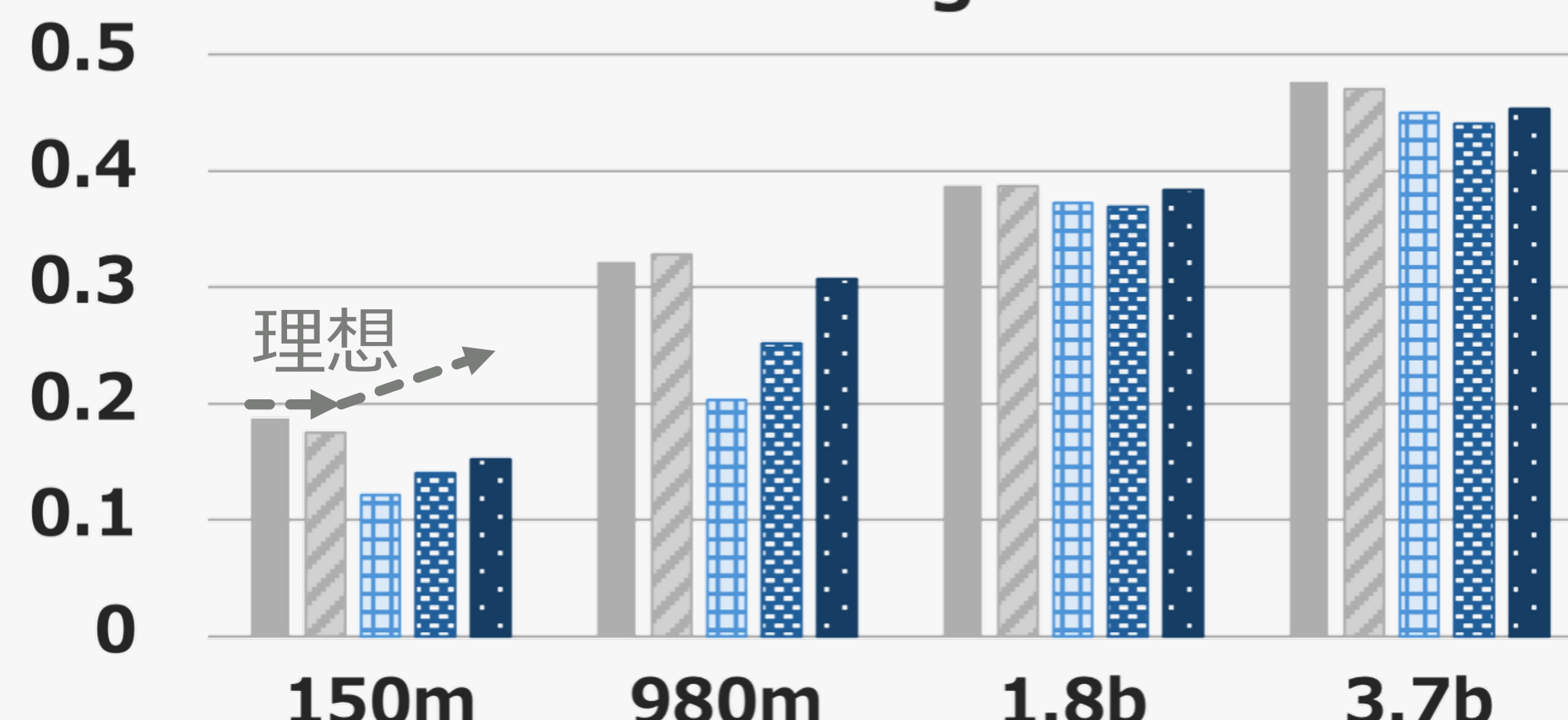


quality: 総合評価



llm-jp-evalでの評価結果

Average



- 生成応答を活用する手法ではベースラインを超えられず
- モデルサイズが大きいほど手法②③の性能の低下が改善
- 生成能力が高い大きいモデルでは性能が向上する可能性

概要

- 学習対象の言語モデル自身が生成かつ選好したデータを活用して事後学習
- 性能向上に効果的なデータ選好手法の検討・検証
- 安全性と有用性の観点で評価
- 安全性は向上したが、有用性は低下した
- 更なる検討・検証が必要

応答選好手法

- 応答候補間のコサイン類似度をもとにしたスコアを利用
- 候補の中で最も中心的な応答を新たな応答にする
cf.) MBR distillation [Wu+ '25]
- 類似度計算には学習対象モデルの埋め込みを利用

今回検証した選好手法

手法①: 元応答を再度利用

手法②: 生成応答から選好

元応答のスコアが下位の場合

手法③: 学習データから除外

手法④: 元応答を利用

	y_i	y_{i1}	y_{i2}	...	y_{in}	平均
y_i	-	0.57	0.51	...	0.48	→ 0.55
y_{i1}	0.57	-	0.68	...	0.50	→ 0.66
y_{i2}	0.51	0.68	-	...	0.42	→ 0.52
...
y_{in}	0.48	0.50	0.42	...	-	→ 0.49

評価方法

- llm-jp-judge [中山+ '25]
- 安全性と品質を5段階評価
- 評価モデル:
gpt-4o-2024-08-06

- llm-jp-eval [Han+ '24]

日本語の大規模言語モデルを複数データセットで横断的に評価 (※ CGIは除いた)

- 生成応答を活用することで大きいモデルでは安全性が向上

- 品質に関して
手法②はベースラインと同等以上
手法③④はベースラインを下回る

- 生成された応答に何らかの有効性
- 品質の観点では、手法③④がうまく機能していない

結論／今後の展望

- 今回検証した手法において性能向上が見られたのは限定的
- モデルサイズをより大きくしたり生成応答数 n を増やしたりした場合の検証
- より効果的なデータ選好法の検討