RAG 학습 방법론

2025. 01. 23

박찬희

RAG 개선 방안

1. Data Preprocessing (Data-Level)

PDF parsing … OCR, vision LLM
chunking … fixed length, semantic embedding
indexing … FAISS, PINECONE

2. Retrieval (Retriever-Level)

Advanced RAG : HyDE, Query Expansion, Reranking

3. Inference (LLM-Level)

Agent
Prompt Engineering
Supervised Fine-Tuning (SFT)
Domain Specific Finetuning (DSF)

RAG 개선 방안

Research Question

컨텍스트 정보를 잘 활용하도록 모델을 **학습**하는 방법은 무엇일까?

Thought Process

- 일반적으로 **SFT**는 **특정 task 수행 (NMT, MRC, SMR)**을 잘하기 위해 이루어짐.
- **컨텍스트 활용 능력**은 일종의 창발능력으로, 다양한 태스크를 잘하기 위한 **수단**이지, 그 자체가 **목적이 아님**.
- 하지만 실제 LLM 사용 사례는 RAG 기반 **open-domain QA**가 상당수이고, 이에 따라 "**RAG를 잘하는 모델** " 에 대한 수요도 늘고 있음.

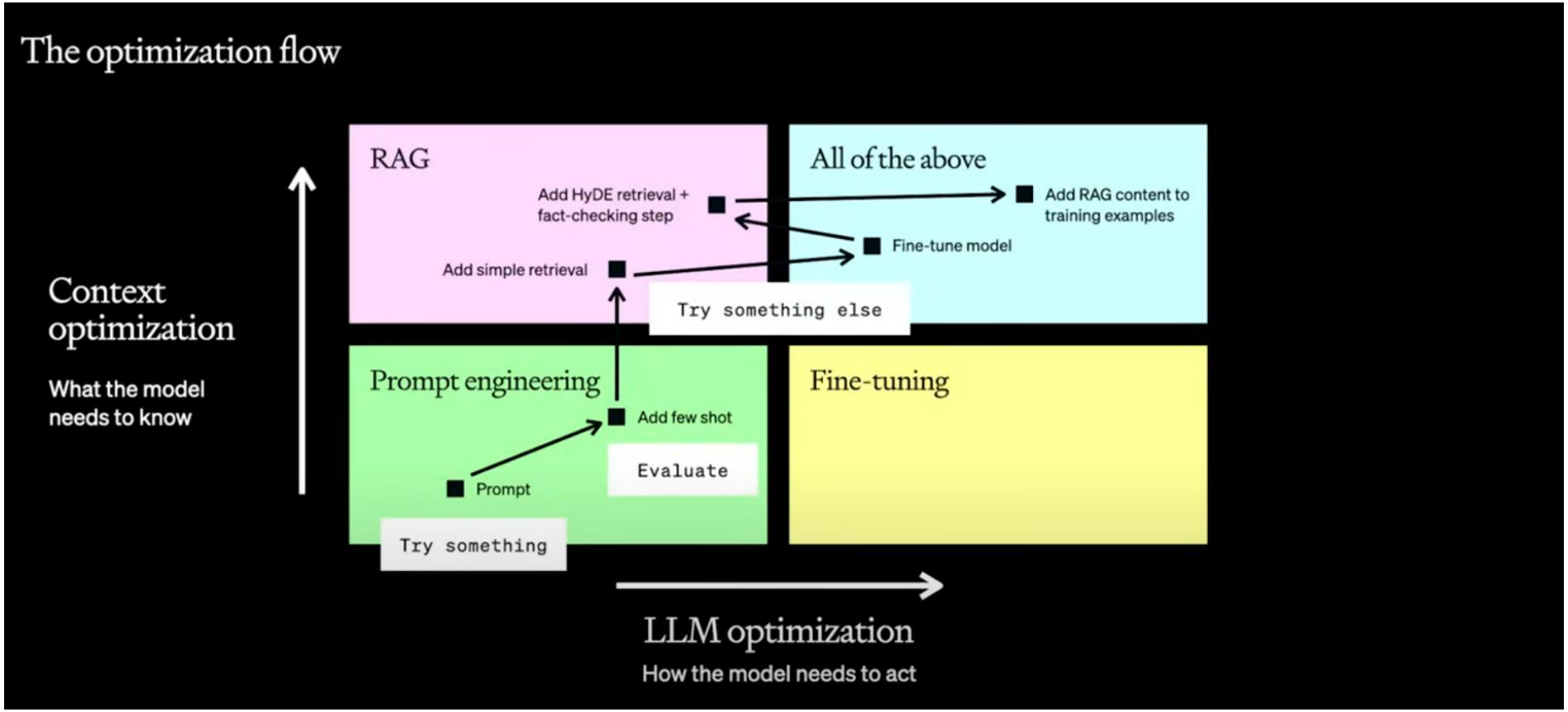
RAG 개선 방안

Domain-specific LLM

- Data governance, security 등 관점에서 **법률, 경제, 의료** 등 도메인 특화 모델을 개발하고자 하는 수요가 증가
- **Pretraining, Instruction Tuning** 등 기존의 방법론은 LLM에 특정 도메인 **지식을 주입**하는 데에는 효과적이지만, 사내 데이터, 표, 이미지 등 **계속해서 업데이트되는 지식**을 **실시간으로 활용**하기 위해서 **RAG**가 적용됨.

RAG 개선 방안

SFT vs RAG?



[A Survey of Techniques for Maximizing LLM Performance]

RAG 개선 방안

Why not both?

도메인 특화 **지식에 대한 학습**과

RAG 기반 답변을 생성하는 **ICL 능력을 향상하는 학습**을

동시에 진행할 수 없을까?

RAFT: Adapting Language Model to Domain Specific RAG

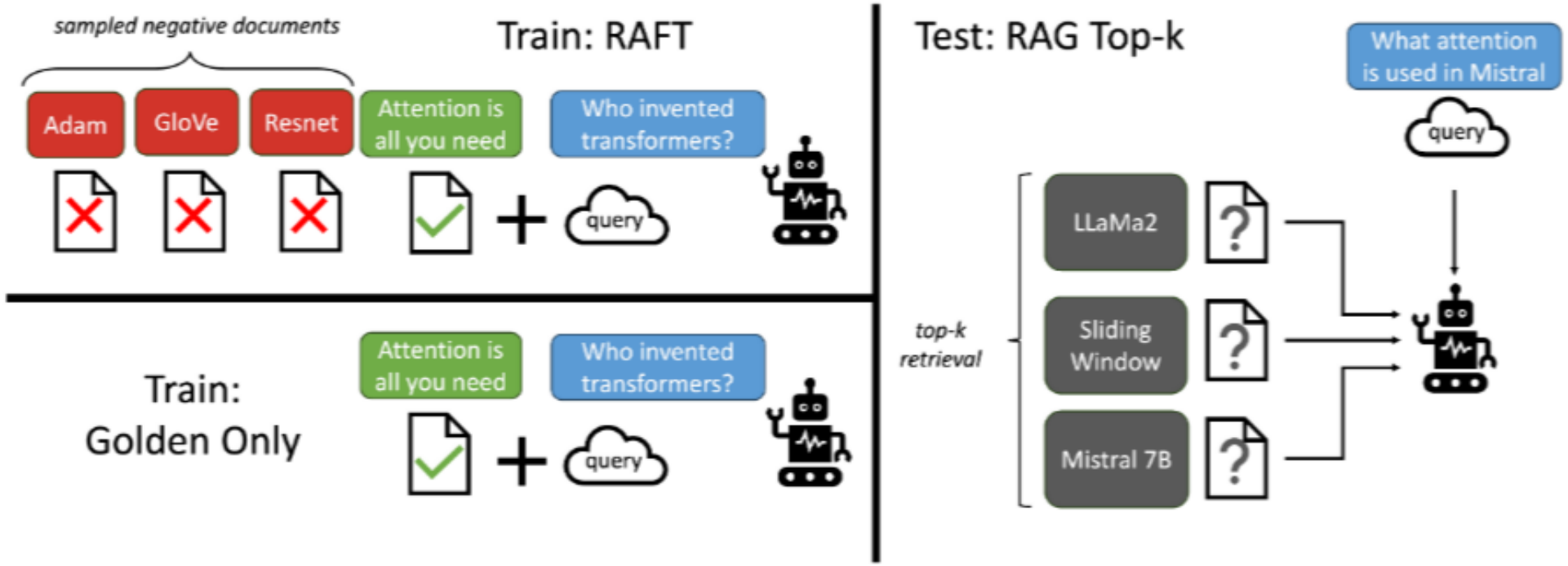
Tianjun Zhang *
Department of Computer Science
UC Berkeley
Berkeley, CA 94720, USA
{tianjunz}@berkeley.edu

Shishir G. Patil, Naman Jain, Sheng Shen
Department of Computer Science
UC Berkeley
Berkeley, CA 94720, USA
{shishirpatil,naman_jain,sheng.s}@berkeley.edu

Matei Zaharia, Ion Stoica, Joseph E. Gonzalez
Department of Computer Science
UC Berkeley
Berkeley, CA 94720, USA
{matei,istoica,jegonzal}@berkeley.edu

Overview

RAFT aims to not only enable models to **learn domain-specific knowledge** through fine-tuning, but also to ensure **robustness against distracting retrieved information.**



Parameterization

Q : question
D_k : a set of documents
    D* : 'golden' document
    D_i : 'distractor' document
A* : CoT answer from 'gold' document (D*)
P : D* + D_k-1의 비율 (hyperparameter)
1-P : D_k로만 구성된 데이터의 비율

$$\mathbf{P}\ \%\ \text{of data: } \mathbf{Q} + \mathbf{D}^* + \mathbf{D}_1 + \mathbf{D}_2 + \ldots + \mathbf{D}_k \rightarrow \mathbf{A}*$$

$$(1 - \mathbf{P})\ \%\ \text{of data: } \mathbf{Q} + \mathbf{D}_1 + \mathbf{D}_2 + \ldots + \mathbf{D}_k \rightarrow \mathbf{A}*$$

Prompt for CoT

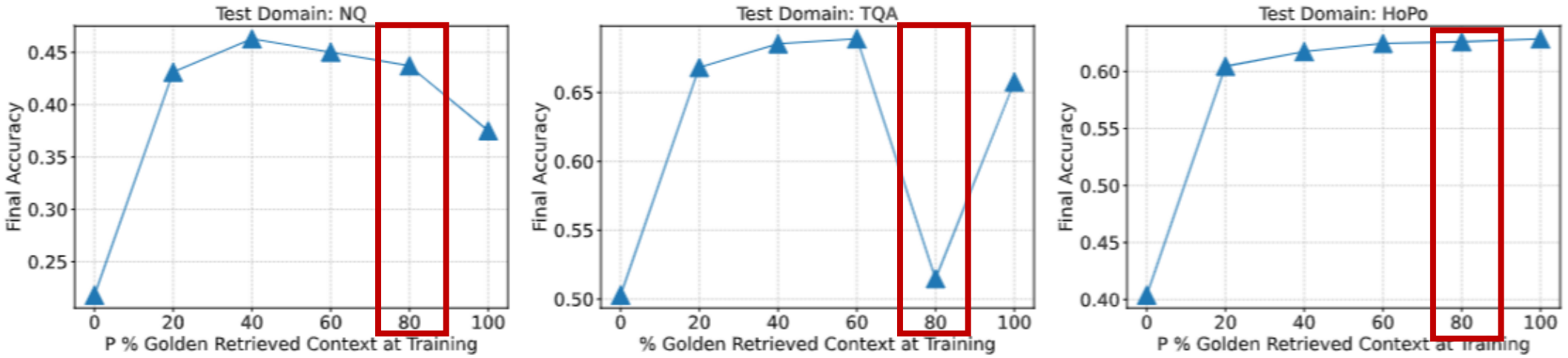
```
Question: The Oberoi family is part of a hotel company that has a head office
in what city?

context: [The Oberoi family is an Indian family that is famous for its
involvement in hotels, namely through The Oberoi Group]...[It is located in
city center of Jakarta, near Mega Kuningan, adjacent to the sister JW Marriott
Hotel. It is operated by The Ritz-Carlton Hotel Company. The complex has two
towers that comprises a hotel and the Airlangga Apartment respectively]...[The
Oberoi Group is a hotel company with its head office in Delhi.]

Instruction: Given the question, context and answer above, provide a logical
reasoning for that answer. Please use the format of: ##Reason: {reason}
##Answer: {answer}.

-------------------------------------------------------------------------------

CoT Answer: ##Reason: The document ##begin_quote## The Oberoi family is an
Indian family that is famous for its involvement in hotels, namely through The
Oberoi Group. ##end_quote## establishes that the Oberoi family is involved in
the Oberoi group, and the document ##begin_quote## The Oberoi Group is a hotel
company with its head office in Delhi. ##end_quote## establishes the head
office of The Oberoi Group. Therefore, the Oberoi family is part of a hotel
company whose head office is in Delhi. ##Answer: Delhi
```

Experiment

	PubMed	HotPot	HuggingFace	Torch Hub	TensorFlow
GPT-3.5 + RAG	71.60	**41.5**	29.08	60.21	65.59
LLaMA2-7B	56.5	0.54	0.22	0	0
LLaMA2-7B + RAG	58.8	0.03	26.43	08.60	43.06
DSF	59.7	6.38	61.06	84.94	86.56
DSF + RAG	71.6	4.41	42.59	82.80	60.29
RAFT (LLaMA2-7B)	**73.30**	35.28	**74.00**	**84.95**	**86.86**
RAFT w.o CoT	68.30	25.62	59.07	**86.56**	83.21

PubMed : **의료** 도메인 QA셋
HotPot : **일반** 도메인 multi-hop QA (NQ, TriviaQA)
HuggingFace, Torch Hub, TensorFlow : **코드** 기반 태스크

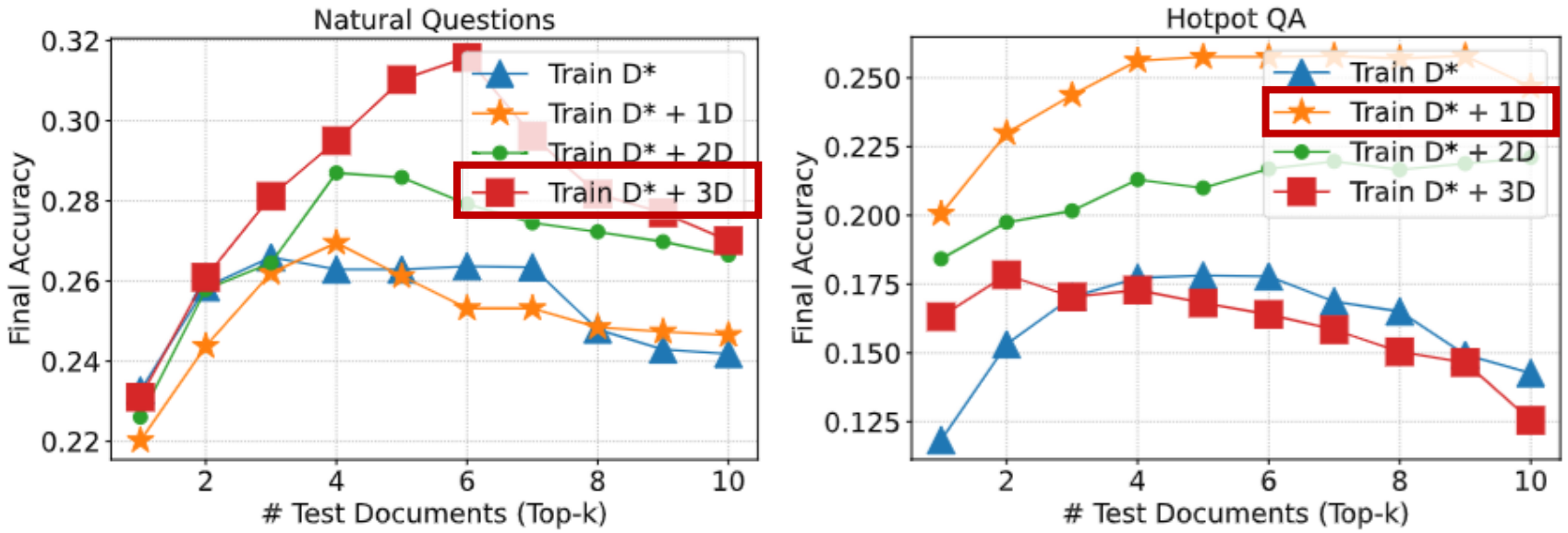
Hyperparameter Setting : P



학습 데이터 중 몇%에 gold passage를 포함시켜야 하는가? (100%가 아니다!)

Training your LLM **without the correct corresponding context** at times **can be beneficial** for the downstream task of answering questions related to the documents. (**k=5**, training 시에 20%는 distractor context로만 구성된 채로 학습함. 즉, **p=0.8**)

Hyperparameter Setting : K



학습 시 distractor를 주는 이유가 무엇인가? (Better Robustness)

실험 결과 distractor context를 함께 준 경우의 최종 성능이 더 좋다!
The inclusion of distractor documents during training indeed **makes the model more resilient to fluctuations** in the number of documents encountered during testing.

정리

장점

- 시의적절한 Novelty
- 새로운 학습의 방향성 제시
- 가능성이 무궁무진함
- 높은 설득력

한계

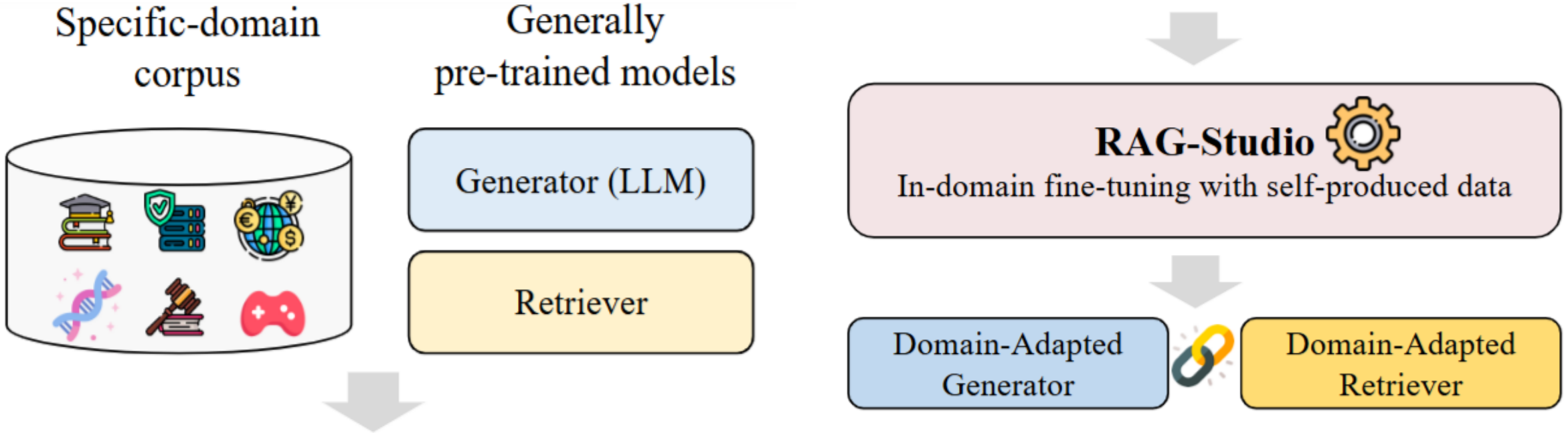
- 어딘가 부족한 설명
- Appendix, github의 부재
- 공개된 모델이 없음

RAG-Studio: Towards In-Domain Adaptation of Retrieval Augmented Generation Through Self-Alignment

Kelong Mao[1], **Zheng Liu**[2*], **Hongjin Qian**[2], **Fengran Mo**[3],
Chenlong Deng[1], **Zhicheng Dou**[1*]
[1]Gaoling School of Artificial Intelligence, Renmin University of China
[2]Beijing Academy of Artificial Intelligence
[3]Université de Montréal, Québec, Canada
{mkl,dou}@ruc.edu.cn, zhengliu1026@gmail.com

Overview

RAG-Studio accepts a specialized **domain corpus**, where it identifies useful domain knowledge and **synthesizes training data** on top of it.



Overview

1. 도메인 코퍼스로부터 QA쌍을 **생성,** 생성된 Query로 관련 정보를 **검색**
2. LLM의 **self-curation**을 통해 검색된 context의 변별력을 판단함
3. 생성된 정보들을 바탕으로 LLM과 retriever를 동시에 **학습**함.

RAFT와의 차별점

- RAFT는 노이즈를 무시하고 답변을 생성하도록 **LLM만** 학습
- RAGStudio는 LLM과 retriever를 **동시에 학습**함. (**Self-alignment**)

Method

1. Raw Data Generation

From $D = \{p_1, ..., p_n\}$

$$q, a = G(\text{Prompt}_{\text{gqa}}(p)),$$

p -> gold passage

$$p_1^c, ..., p_K^c = R(q, D).$$

Each sample $s = (q, a, p, \{p_1^c, ..., p_K^c\})$

(본 실험에서 k=3)

Method

2. Data Curation for Generator

Using CoT prompting, each context *p* is labeled as

helpful, *irrelevant*, *misleading*

$$e', a' = G(\text{Prompt}_{\text{rag}}(q, C)), \longrightarrow \begin{aligned} C_1 &= \{p\} \cup \{p_1^c, \ldots, p_K^c\} \\ C_2 &= \{p_1^c, \ldots, p_K^c\} \end{aligned}$$

자신이 생성한 답변에 대하여 q, p, a를 주고 판단
처음부터 맞춘 샘플은 제외, 틀렸다면 답을 주고 rationale 생성

=> (e' a')는 "잘못된 사례"를 수집하기 위함

$$e = G(\text{Prompt}_{\text{rationale}}(q, C, a)).$$

Method

2. Data Curation for Generator

Rationale e is filtered :

- For $C = C_1$, if the ground-truth passage p is not labeled as *helpful* in e, the sample is discarded.
- For $C = C_2$, if more than two passages are labeled as *helpful* in e, the sample is discarded.

최종 데이터는 다음과 같은 triplet : $(q + C, e + a, e' + a')$

입력문 | 선호 답변 | 비선호 답변

Method

3. Data Curation for Retriever

a′ is correct: *helpful* -> **positive**
misleading -> **hard negative**

a′ is incorrect: *helpful* -> **hard negative**

Original passage *p* is always a **positive** sample

Method

4. Finetuning

Generator : ORPO

x : (q+C)
Y^w : (e+a)
Y^l : (e′+a′)

$$\mathcal{L}_{\mathrm{G}} = -\lambda \log \sigma(\log \frac{P_\theta(y^w|x)(1-P_\theta(y^l|x))}{P_\theta(y^l|x)(1-P_\theta(y^w|x))}) - \frac{1}{m}\sum_{t=1}^{m} \log P_\theta(y_t^w|x, y_{<t}^w),$$

Retriever : Contrastive Ranking

p^+ : positive samples
p^- : negative samples

$$\mathcal{L}_{\mathrm{C}} = -\log \frac{\phi(x, p^+)}{\phi(x, p^+) + \sum_{p^- \in P^-} \phi(x, p^-)},$$

Experiment

Dataset

Statistic	Biomedical	Finance	Law	Computing	TriviaQA (wiki)
#Questions (train)	26,877	8,245	3,770	3,668	61,888
#Questions (test)	4,743	1,455	666	648	7,993
#Passages	287K	89K	42K	41K	2M

앞의 4개 데이터셋은 MS MARCO에서 발췌한 **벤치마크 데이터**
실험에 사용한 RAGStudio 데이터 합성에는 **Llama-3-8B-Instruct**와 **BGE**로 각
각 **3000개**씩 합성한 sample을 사용함. (Retrieval top-K는 3)

Experiment

Metric

ROUGE-L : domain-specific QA
EM : TriviaQA
GPT-4 : answer accuracy 측정

Baselines

(1) **GPT-3.5** : GPT + 프롬프팅
(2) **Prompt** : llama + 프롬프팅
(3) **DSF** : MS MARCO, TriqiaQA의 만들어진 학습셋 사용.
(4) **RAFT** : P=0.6으로 재현
(5) **RAGStudio** : **LoRA** 학습

(관건은, DSF (human-labeled data)보다 좋은가?

Experiment

Result

Method	FT	Synthetic Data	Biomedical		Finance		Law		Computing		TriviaQA (wiki)	
			R-L	Acc	R-L	Acc	R-L	Acc	R-L	Acc	EM	Acc
without Retrieval												
GPT-3.5	✗	✗	19.5	43.1	16.9	34.2	20.5	51.4	19.6	42.9	62.9	60.0
Prompt	✗	✗	22.2	45.4	18.4	35.7	19.7	50.1	21.6	46.3	64.8	60.5
DSF	✔	✗	23.3	33.7	23.2	31.8	23.1	35.9	22.1	34.4	59.7	56.7
with Retrieval												
GPT-3.5	✗	✗	31.2	55.0	36.5	52.0	35.0	57.5	31.8	54.6	69.5	69.5
Prompt	✗	✗	31.9	54.7	37.7	52.2	34.1	56.2	34.9	56.9	69.3	69.2
DSF	✔	✗	34.8	57.5	40.0	55.4	37.4	60.4	44.2	57.9	**70.2**	**69.9**
RAFT	✔	✔	34.3	55.9	36.5	51.7	36.7	59.8	45.4	59.2	69.4	69.6
RAG-Studio	✔	✔	**37.5**	**59.2**	**41.8**	**57.6**	**40.0**	**63.5**	**47.1**	**60.3**	70.0	**69.9**

Experiment

Result

(1) TriviaQA를 제외하곤 RAGStudio의 성능이 가장 좋다.

(2) 도메인 특화 데이터에서의 RAGStudio 상승폭이 크다.

(3) 검색 없이 FT만 한 경우 **(DSF)** 성능이 오히려 **떨어진다.** (일반화 능력 감소)

Experiment

Ablation

Method	$\overline{\Delta}$	Bio.	Fin.	Law	Comp.	TQA
Prompt	-4.3	54.7	52.2	56.2	56.9	69.2
DSF	-1.9	57.5	55.4	60.4	57.9	69.9
SFT	-1.2	58.5	56.7	62.2	57.9	69.3
w/o G-FT	-3.9	54.9	53.4	57.4	56.5	68.8
w/o R-FT	-0.8	58.7	57.2	62.4	58.8	69.6
RAG-Studio	0	59.2	57.6	63.5	60.3	69.9

SFT : (x+C) -> (e+y)로만 학습한 경우
DPO도 실험했으나, 성능 일관성이 떨어져 제외

Experiment

Effects of CoT

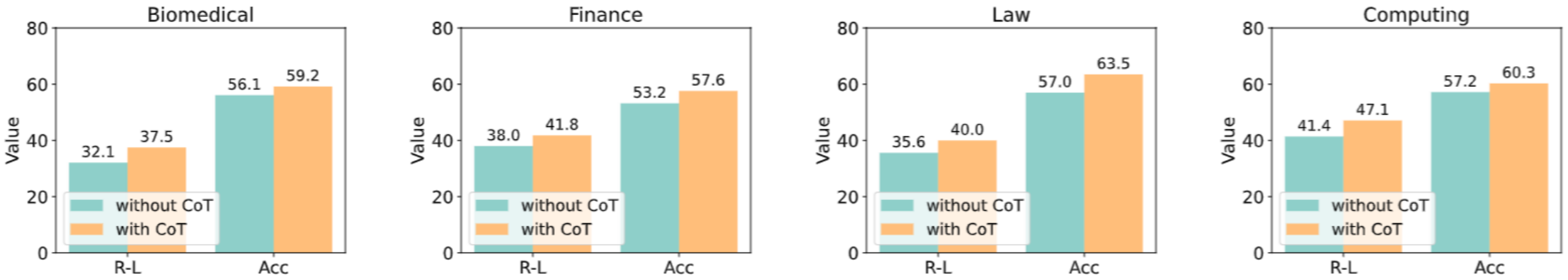


Figure 3: Comparison of RAG-Studio performance with and without our tailored chain-of-thought for fine-tuning.

Generator : (e+a)에서 (a)만 학습
Retriever : hard negative 없이 in-batch negative로만 학습

Experiment

In-domain Retrieval

Retriever	Biomedical		Finance		Law		Computing		TriviaQA (wiki)	
	R@5	NDCG@5	R@5	NDCG@5	R@5	NDCG@5	R@5	NDCG@5	R@5	NDCG@5
Before FT	50.8	36.1	66.6	49.4	77.3	58.6	77.5	56.5	23.1	34.1
After FT	51.4	37.9	66.2	51.9	77.1	59.5	78.3	59.0	22.5	34.4

Table 4: Comparison of retrieval performance before and after in-domain fine-tuning.

Retriever 역시 FT 후 전반적인 성능이 향상됨

Experiment

Data Quantity -> 다다익선 but saturate

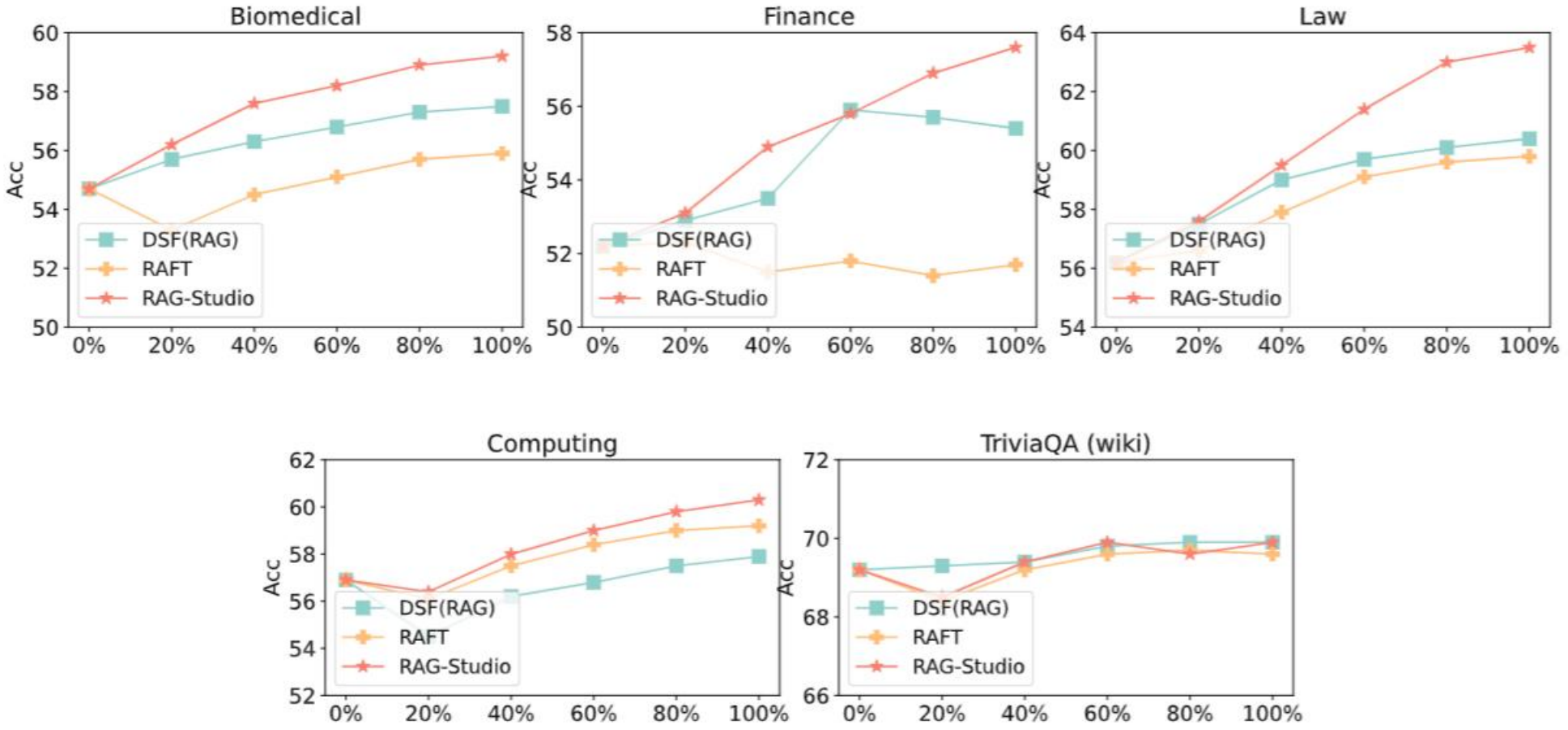


Figure 4: Performance of different methods using different percentages of training data.

Experiment

Data Quality

Method	Biomedical		Finance		Law		Computing		TriviaQA (wiki)	
	R-L	Acc	R-L	Acc	R-L	Acc	R-L	Acc	EM	Acc
RAG-Studio	37.5	59.2	41.8	57.6	40.0	63.5	47.1	60.3	70.0	69.9
Using Raw Data	35.0	56.5	37.6	53.1	38.0	60.8	44.9	58.8	69.2	69.5

Table 5: Performance comparisons of fine-tuning on data with and without filtering.

Filtering으로 약 5% 내외의 성능 향상

Experiment

Different Model Setups

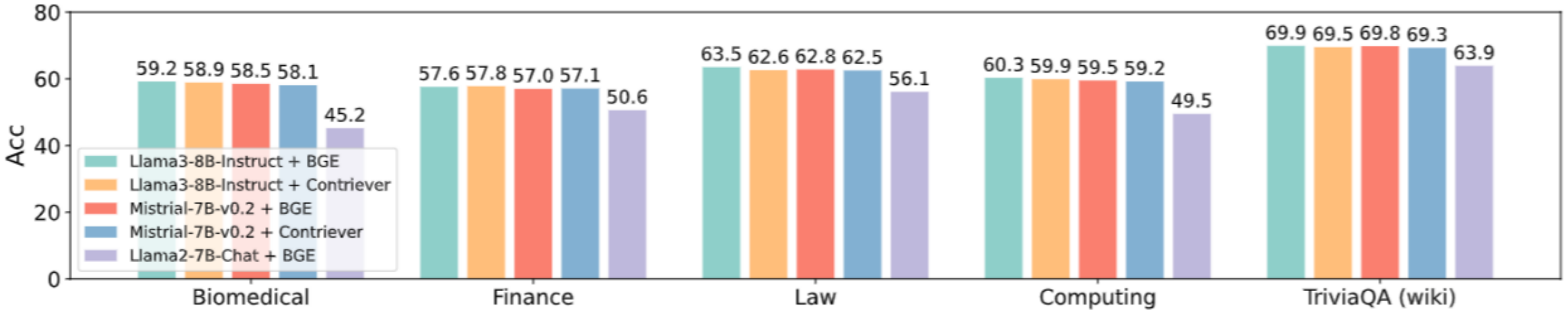


Figure 5: Performance of RAG-Studio variations based on different initial generator and retriever combinations.

Weaker LLM -> lower data quality

정리

장점

- 자동화된 3-step 학습 프레임워크
- 구체적인 실험 설계
- 학습 데이터 필터링과 검증

한계

- LORA가 아니라 full FT했다면 성능이 더 좋았을 텐데
- 3000개라는 다소 아쉬운 데이터 수량

In-Domain RAG 모델 개발 가능성

Q. 특정 분야에서 GPT-4o를 능가하는 Llama 3 8B 모델을 개발할 수 있을까?

RAG-Studio는 3000개의 자체 제작한 합성 데이터를 LORA 학습한 결과만으로 꽤 유의미한 성능 향상을 보임.

RAG 파이프라인 구축 시 DB 수집이 가장 우선시되는데, 이렇게 수집한 문서를 바탕으로 In-domain adaptability를 강화하는 방향으로 모델의 잠재력을 끌어올릴 수 있을 것으로 기대.

Thank you