南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

组会汇报

2025年1月6日

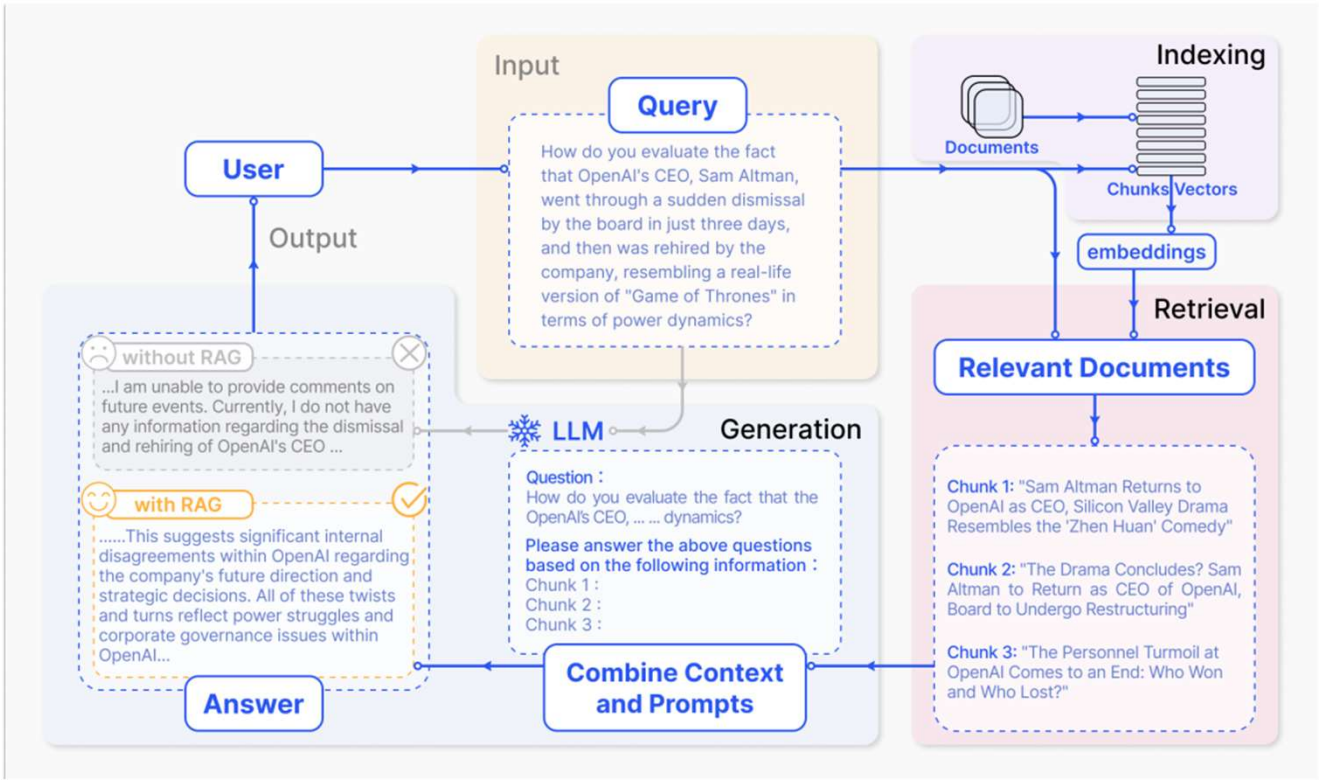
RAFT: Adapting Language Model to Domain Specific RAG

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

RAG（Retrieval-Augmented Generation）



RAG（检索增强生成）过程在问答任务中的一个典型实例主要包括以下三个步骤：

索引：将文档分成小块，编码为向量，并存储在向量数据库中。

检索：根据语义相似度检索与问题最相关的前k个文档块。

生成：将原始问题与检索到的文档块一起输入大语言模型（LLM），生成最终答案。

RAG的优势与挑战

动态获取最新信息：
RAG允许模型从外部检索文档最新的信息，因此能够回答超出模型预训练知识范围的问题，适应性更强。

提升回答准确性：
通过检索相关文档并生成答案，RAG能够提供更加准确、基于事实的答案，特别是在需要引用具体文档时表现更好。

领域适应性强：
RAG适用于特定领域的问答任务（如法律、医学等），可以灵活调整知识库，适应不同的领域需求。

“幻觉”现象：
生成模型固有的问题，缺乏充足信息支持时，模型倾向于生成看似合理实则偏离事实的内容。

索引数据质量：
RAG的性能依赖于检索系统的数据质量。如果索引中的文档或数据不准确、过时或与查询意图不匹配，会直接影响到检索的准确性。

检索查询构建：
生成的检索查询可能没有充分捕获用户的真实意图或者与索引中的信息对齐，导致输入检索不准确。

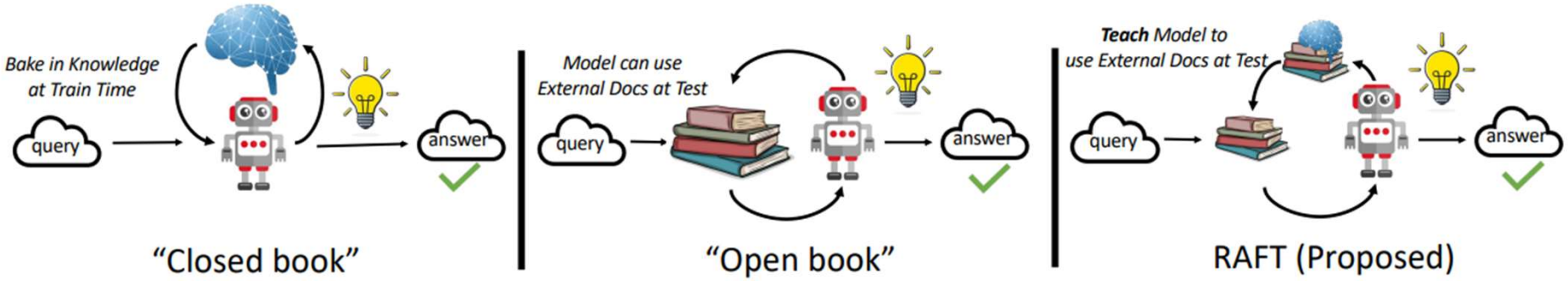
SFT（Supervised Fine-Tuning）

SFT是微调的一种形式，强调在有监督的环境下进行。SFT是一种简单的微调方法，它使用带有正确答案的数据集来继续训练一个预训练的模型。这种方法依赖于大量的标注数据，即每个输入都有一个预先定义的正确输出。 微调的目的是使模型更好地适应特定的任务或领域，比如特定类型的语言理解或生成任务。SFT通常不涉及复杂的策略或奖励函数，只是简单地最小化预测输出和真实输出之间的差异。

Motivation

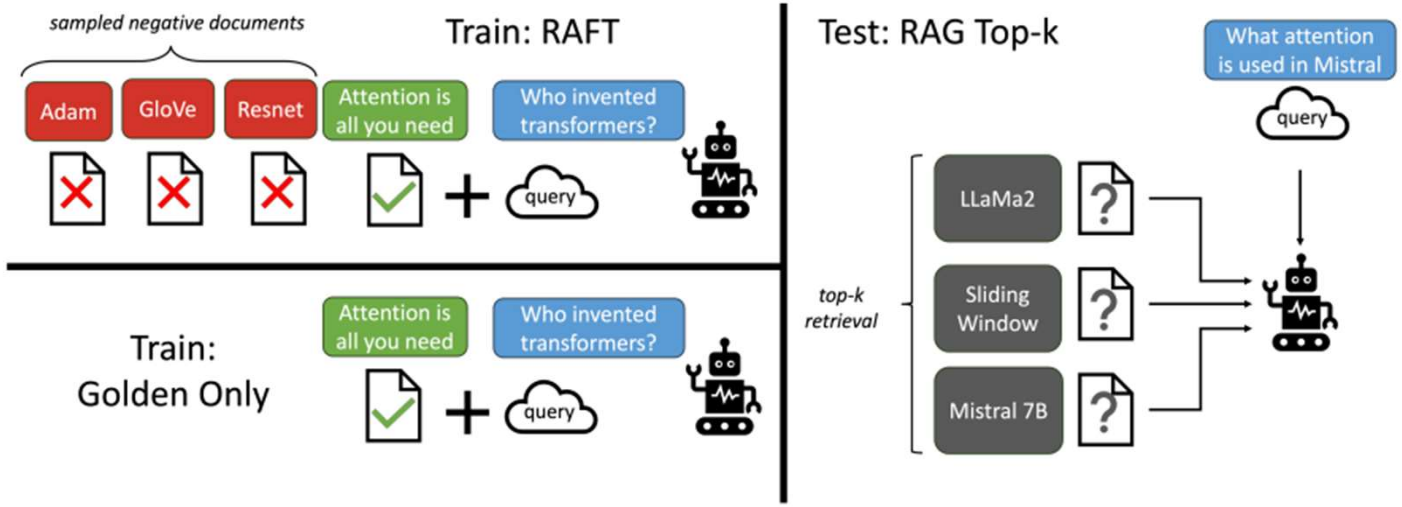
- 大模型引入新知识（例如紧急新闻或私有领域知识）要么通过RAG，要么微调，但如何最佳地让模型获得这些新知识还值得探索。
- 为了改善LLMs在特定领域中的性能，特别是在开放书本式问答（"open-book" questioning）任务中，本文提出了RAFT（Retrieval Augmented FineTuning）方法，尝试将SFT，RAG，COT进行结合，提升特定领域QA的效果。

Methods——How best to prepare for an Open Book Exam?



纯监督的微调（SFT）：种方法通过直接"记忆"输入文档或者回答练习问题（不引用文档）来"学习"，类似于考试前直接背诵或者通过做练习题来准备，但没有引用外部文档。
纯检索增强生成（RAG）：这种方法相当于开卷考试，但是没有提前学习。即在回答时可以引用文档，但没有充分利用固定领域设置和提前接触到测试文档的学习机会。
监督微调和检索增强生成结合（RAFT）：结合了监督式微调（SFT）和检索增强生成（RAG），通过引用文档中的相关段落来生成答案，同时在模拟的不完美检索环境中进行训练，从而有效地为开卷考试设置做准备。

Methods



训练数据中每条数据包含一个问题（Q）、一组文档（Dk）以及从其中一个文档（D*）生成的链式思维风格答案（A*）。该方法区分了两类文档：“黄金”文档（D*），即可以从中推导出问题答案的文档，和“干扰”文档（Di），这些文档不包含与答案相关的信息。

对于数据集中P%的问题，该方法保留黄金文档（d_i^*）以及干扰文档（d_{k-1}）。对于（1-P）%的问题，我们仅包括干扰文档，不包括黄金文档。接着，我们使用标准的监督训练（SFT）技术微调语言模型，训练模型从提供的文档和问题中生成答案。

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

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

Methods

```
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
```

提升训练质量的关键因素之一是生成推理过程，例如链式思维（Chain-of-Thought，**CoT**），用于解释提供的答案。RAFT的方法类似：本文证明了创建完整的推理链，并且清晰引用来源可以提高模型回答问题的准确性。

左图展示了这一设置。以这种方式生成训练数据，涉及向模型提供问题、上下文和已验证的答案，然后要求其形成适当引用原始上下文的推理链。

实验证明，添加详细的推理段落有助于提升模型的表现。

Evaluation

	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**

无论是否使用RAG，LLaMA-7B模型在回答问题时表现不佳，因为其回答风格与标准答案不匹配。通过应用领域特定的微调，模型的表现得到了显著提升。这一过程使得模型能够学习并采用合适的回答风格。然而，将RAG引入领域特定微调（DSF）的模型并不总是带来更好的结果，这可能表明模型在处理上下文和提取有用信息方面存在不足。通过使用RAFT方法，不仅能够训练模型使其回答风格与要求相匹配，还提升了其文档处理能力。因此，RAFT方法在所有任务中都表现优于其他模型。

Evaluation——Effect of CoT

	PubMed	HotpotQA	HuggingFace	Torch Hub	TensorFlow
RAFT w.o CoT	68.30	25.62	59.07	**86.56**	83.21
RAFT	**73.30**	**35.28**	**74.00**	84.95	**86.86**

模型类型	描述
RAFT w.o CoT	应用了 RAFT 方法但没有使用 Chain-of-Thought 的模型
RAFT	应用了 RAFT 方法并结合了 Chain-of-Thought 的模型

CoT方法显著提高了模型在特定领域内的性能，特别是在处理需要详细推理和解释的任务时。通过生成详细的推理链并引用上下文，模型能够更清晰地解释其答案，从而提高了答案的准确性和可解释性。这表明CoT是提高模型性能的一个重要组成部分，尤其是在需要模型进行深入理解和推理的复杂任务中。

Evaluation——always golden context for RAG?

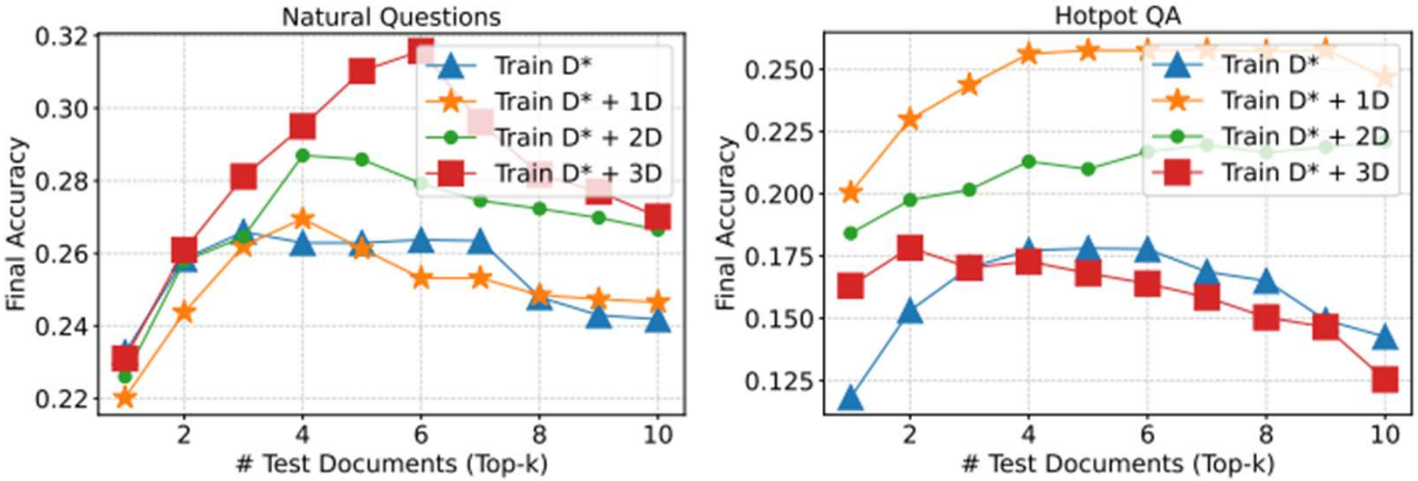
该图展示了对超参数P%的研究，该参数表示训练实例中应包含黄金文档的比例。结果显示，不同数据集的最优比例各不相同，P%的范围在40%、60%到100%之间。这意味着，在某些情况下，在训练LLM时偶尔不提供正确的上下文，可能有助于基于文档的下游问答任务。在训练设置中，四个干扰文档与黄金文档一起被使用，测试时也保持这种格式，提供黄金文档与四个干扰文档。研究结果表明，对于领域特定的RAG任务，在训练数据中包含一定比例没有黄金文档的上下文是有益的。

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

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

Evaluation——Making Model Robust to top-K RAG

不同曲线代表着在训练数据中加入不同数量的干扰文档对模型表现的影响



训练数据配比的重要性，并非所有情况都适合使用相同数量的干扰文档进行训练。例如，在某些数据集上，训练时使用一个“正样本”文档加上三个干扰文档（D* + 3D）可能会获得最佳性能，而在其他数据集上，可能需要调整这个比例。

在训练过程中模拟真实检索环境的重要性。通过在训练中包含不同数量的干扰文档，模型可以学习在测试时如何处理检索器可能返回的不确定数量的文档，包括正确和错误的信息，在有些任务上可以显著提升效果。

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

Conclusion

这篇论文介绍了一种新的训练方法——检索增强型微调（RAFT：Retrieval Augmented Fine Tuning），旨在改善领域特定的RAG（Retrieval-Augmented Generation）条件下，大型语言模型（LLMs）对问题的回答能力。

简单来说，RAFT将RAG+SFT+COT有机的结合起来，在多项问答测试集上，利用LLama2 7B的模型，就取得了比ChatGPT3.5+RAG结合的方法。具体来说，在RAFT中，面对一个问题及一组检索到的文档时，模型被训练为忽略那些不相关的知识，同时挖掘其中与问题的相关知识的来提升问答效果。同时其结合链式思考COT策略提升推理能力。其在PubMed、HotpotQA、Hugging Face、Torch Hub、TensorFlow Hub等问答数据集上，取得了非常不错的效果。

南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Thanks !

2025年1月6日