Adapting Large Language Models for the Dynamic World

Presenter: Zixuan Ke
https://vincent950129.github.io/

LLMs in A Fixed World?

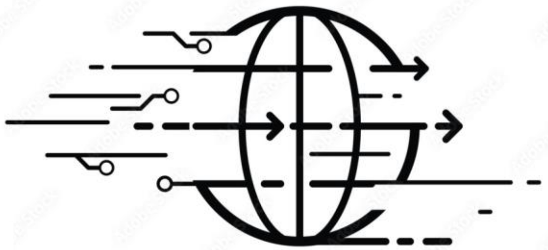
Packed with knowledge and excels in many tasks

The world is **fixed** (i.i.d)

Once trained, LLMs are fixed

LLMs in A Dynamic World!

Packed with knowledge and excels in many tasks



The world is **ever-changing**

Emerging domains/events/topics /information

LLMs in A Dynamic World

Packed with knowledge and excels in many tasks

How to adapt LLMs for the **dynamic world?**

Emerging domains/events/topics /information

LLMs in A Dynamic World: Plan

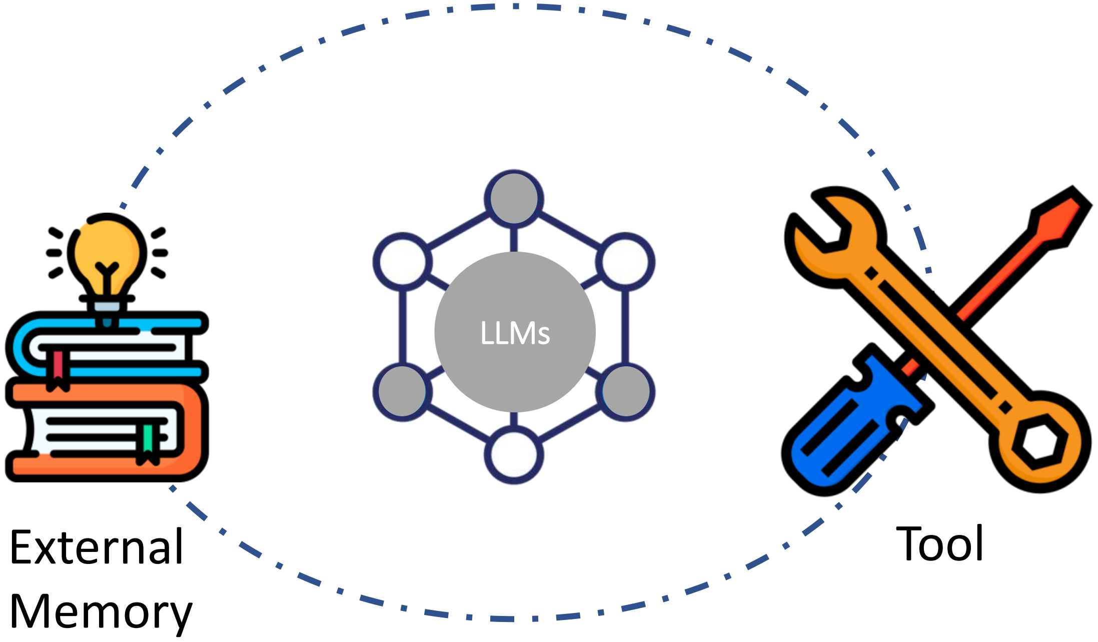
How to adapt LLMs for the **dynamic world?**

- **Black-box LLM**: Retrieved-augmented Generation (RAG)
- **White-box LLM:** Continual Pre-training
- Future Work

Bridging the Preference Gap between Retrievers and LLMs, Ke et al, arXiv 2024
Continual Pre-training of Language Models, Ke et al, ICLR 2023
Adapting a Language Model While Preserving its General Knowledge, Ke et al, EMNLP 2022

LLMs in A Dynamic World

How to adapt LLMs for the **dynamic world?**



Main idea: integrating fresh, external information to the LLMs **without** retraining the LLMs (no need to worry about the LLMs' parameters)

Retrieval-augmented Generation (RAG)

Retrievers

Retrieve task-relevant information,
pack it into the context of the LLM

Existing work fine-tunes retrievers or LLMs or both to improve downstream tasks

A general belief: ranking is the most important, as humans read from top to bottom

However, LLMs may exhibit preferences **different from humans** and yield sub-optimal predictions using the retrieved information

Retrieval-augmented Generation

Retrievers

Retrieve task-relevant information,
pack it into the context of the LLM

retrieves **user-friendly** information

Fundamental Problem:
There could be **preference gap** between the two
(e.g., ranking, selection, repetition)

requires **LLM-friendly** information

Bridging the Preference Gap between Retrievers and LLMs, Ke et al, arXiv 2024

Retrieval-augmented Generation

Retrievers

retrieves **user-friendly** information

Retrieve task-relevant information,
pack it into the context of the LLM

Our focus:

- ❑ Establish the preference gap
- ❑ Propose an approach to bridge the gap

requires **LLM-friendly** information

Our work focus on retrieving passages

Dataset

	#Training	#Val.	#Test	Avg. #Tokens
NQ	79,168	8,757	3,610	517.82
HotpotQA	68,659	5,600	5,600	564.83
Email	13,305	764	1,227	173.85
Book	20,789	41,331	41,331	124.52

Context length < maximum length

Question Answering (NQ and HotpotQA):
Candidate passages are retrieved from WikiPedia Pages

Personalized Generation (Emails and Books):
Candidate passages are retrieved from reviews/emails authored by the same user in the past

Query

Instruction: Finish the passage in the user voice
Review title: Perfect solution for long-range planning!
Review product: 2018 - 2022 artwork five-year planner…
Review start: Wow! I've been searching for something like this and was so pleased when it came in! the

Target

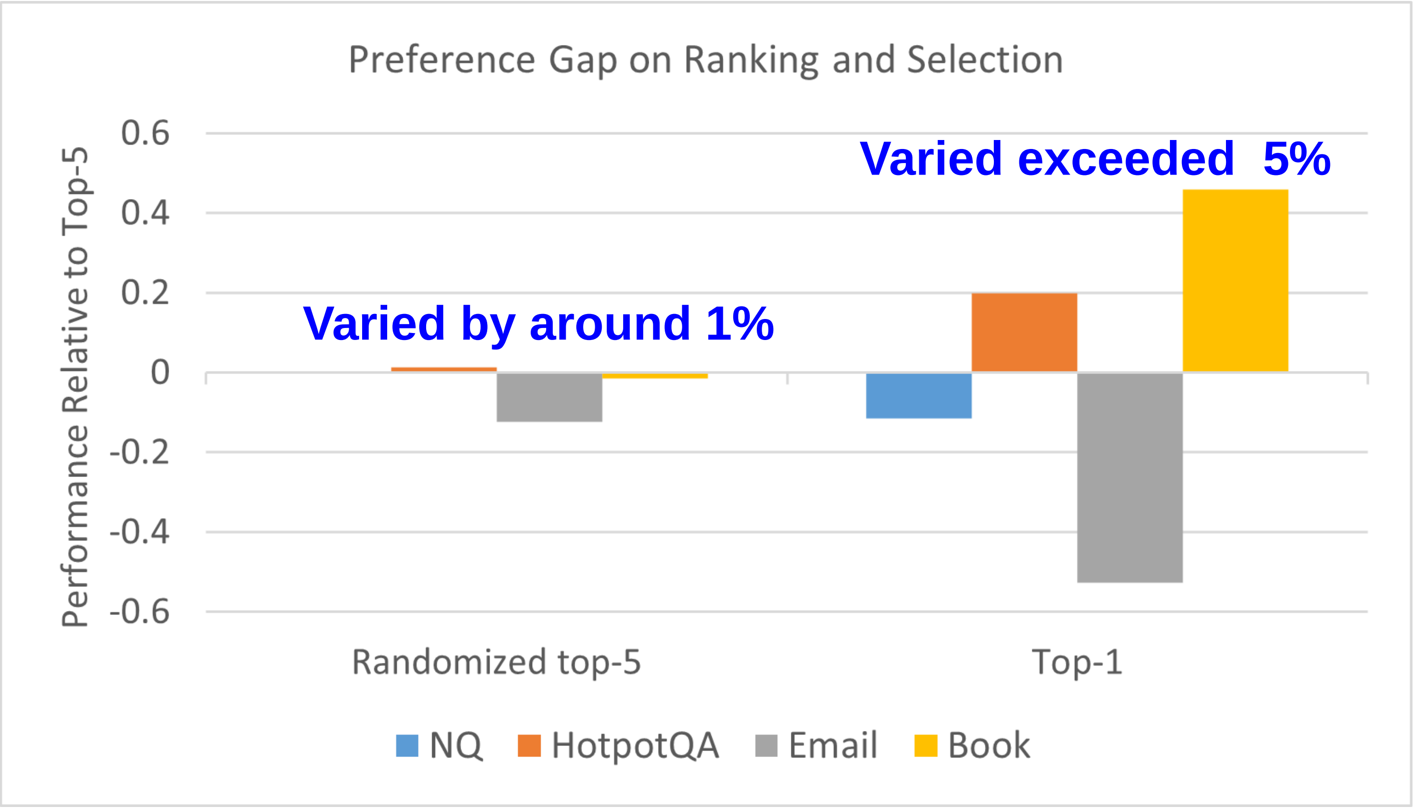
Remaining part: 2-page-per-month style works. The blocks on the calendar are big enough to write quite a bit….

Preference Gap

Ranking: reads sequentially and order is crucial
Selection: can ignore irrelevant

Ranking: order does not impact much
Selection: significantly impact (either positively or negatively)
……(potentially more, e.g., repetition)

The general belief that ranking is most important DOES NOT hold for LLMs!

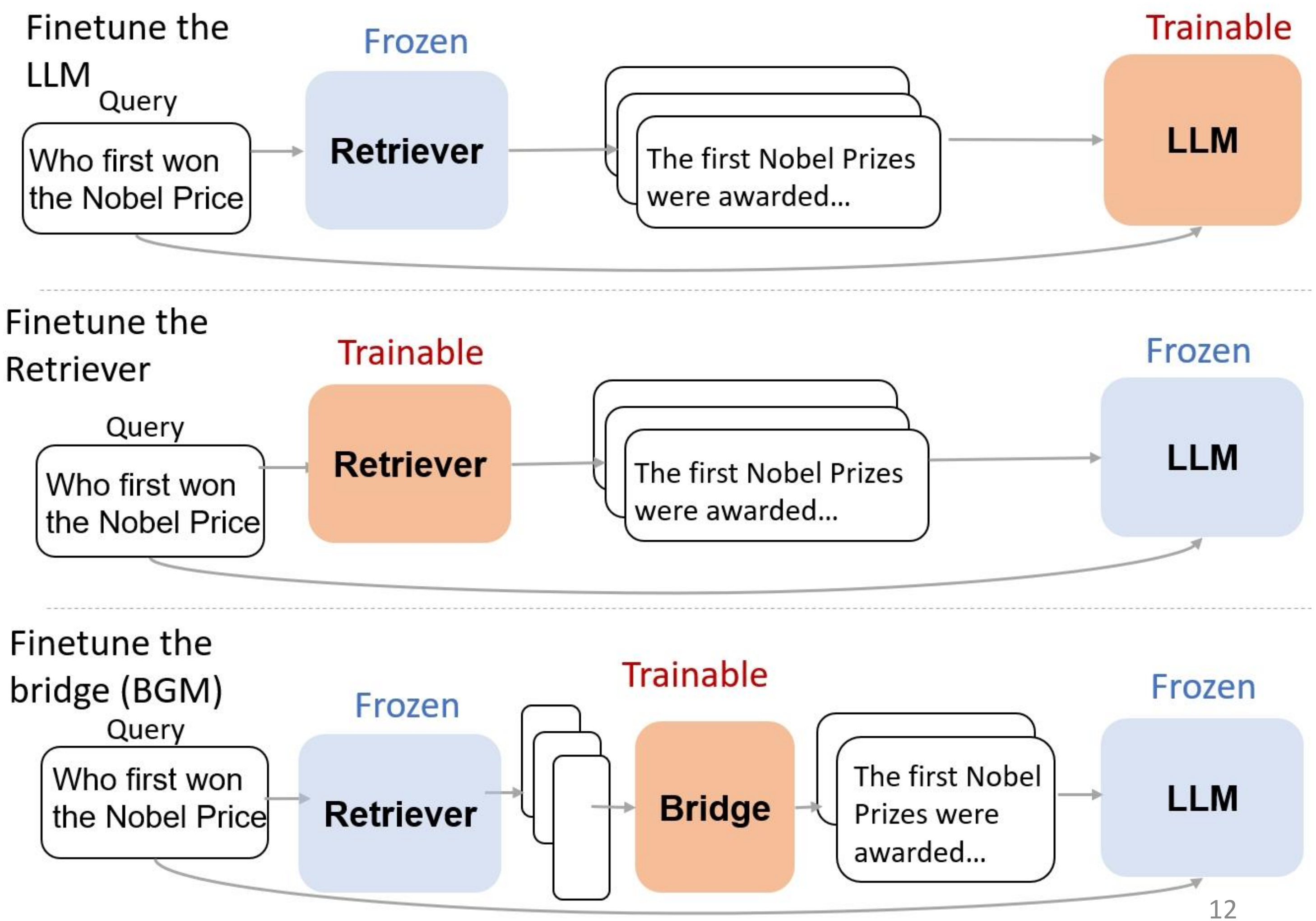


This is a crucial insight as it **confirms** the preference gap and highlights the importance of **bridging this preference gap** to enhance RAG.

Bridging the Gap: Approach

Bridge model

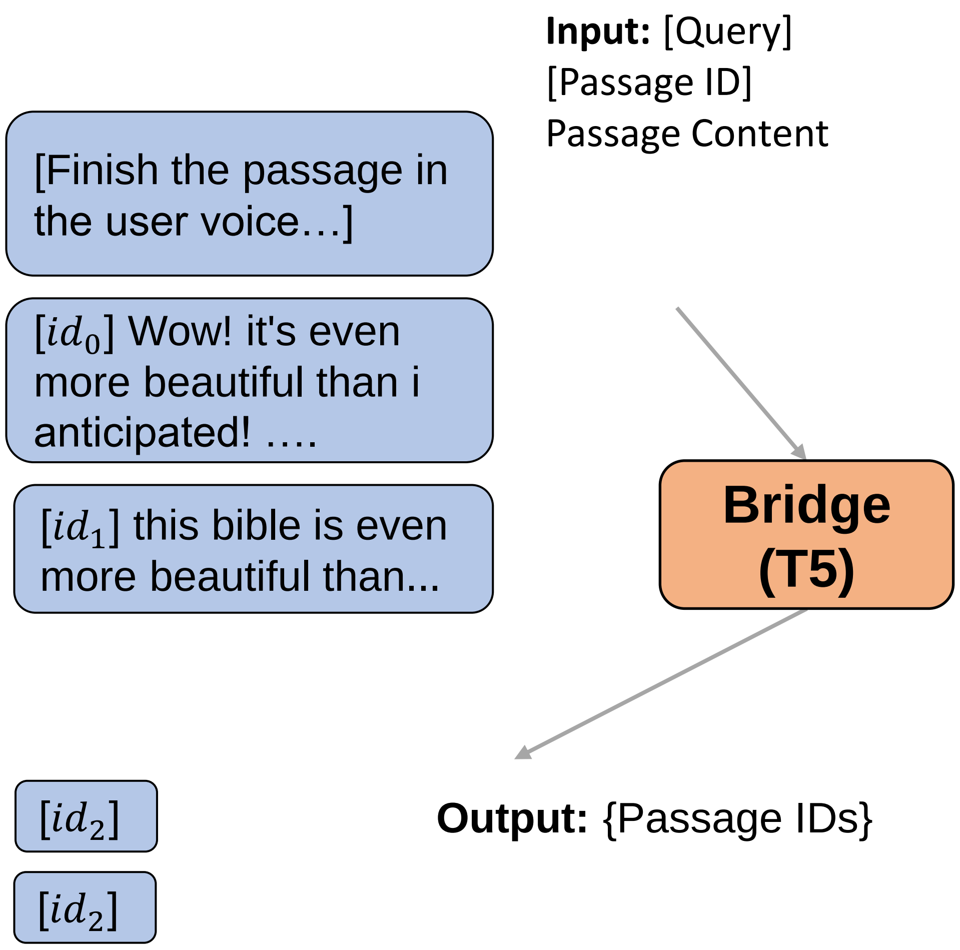
- ❑ **Fix** the Retriever and the LLM and train an intermediate **bridge model**
  - ❑ LLMs are often only available as black-box **APIs** and fine-tuning is not an option
  - ❑ Retrievers **only consider reranking**, not applicable to other possible preference gap



Bridging the Gap: Approach

Seq2seq Format

- ❑ Not only rerank, but also dynamically select passages for each query
- ❑ Potentially employ more advanced strategies like repetition



Bridging the Gap: Approach

Typical RAG

- ❑ No ground truth relevance label for what should be retrieved
- ❑ But only ground truth label for the downstream tasks

Existing Approaches: Supervised learning

- ❑ Use the supervision provided by the LLM, such as the perplexity of downstream tasks
  - ❑ E.g., Feed candidate passage into LLMs and use the perplexity as relevance score
  - ❑ Only Point-wise suspension!

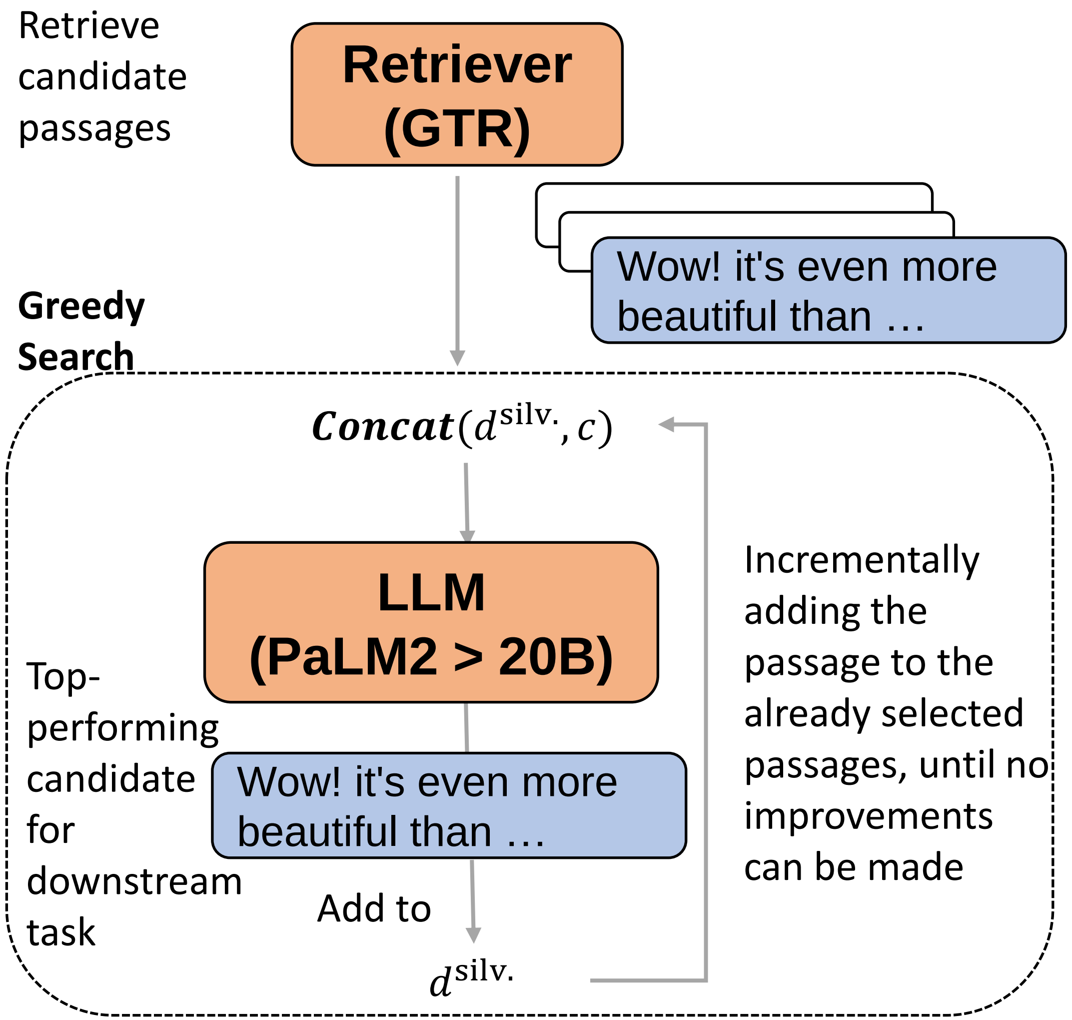
However

- ❑ Sequential supervision is missing or sparse
  - ❑ Nearly impossible to feed all possible retrieved sequences into the LLM to obtain supervision
- ❑ Rely on intermediate relevance label
  - ❑ Not end-to-end training on the downstream tasks

Bridging the Gap: Approach

BGM: Supervised Learning + Reinforcement Learning

- ❑ Supervised Learning
  - ❑ Synthesizing silver passage sequence based on greedy search
    - ❑ We select only the useful passages by incrementally selects the next passage that maximized the downstream task performance

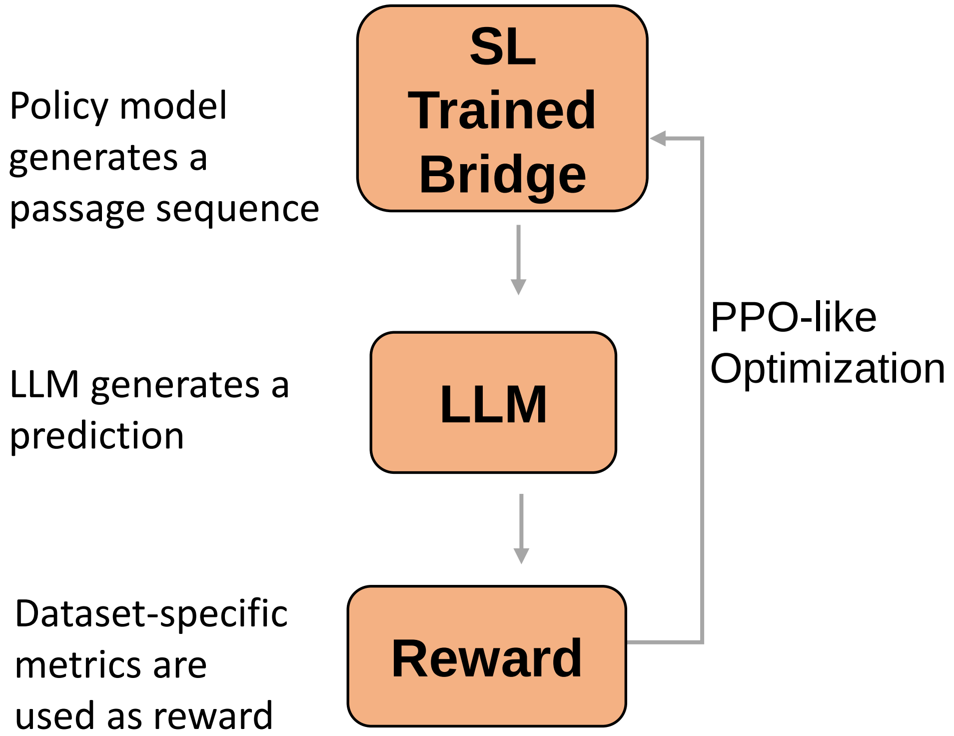


GTR: Large dual encoders are generalizable retrievers, Ni et al, EMNLP 2022

Bridging the Gap: Approach

BGM: Supervised Learning + Reinforcement Learning

- ❑ Reinforcement learning
  - ❑ Downstream task performance as reward, passage IDs as action space, bridge model as policy model
    - ❑ Much more supervision (recall that we only consider permutation or deletions in the silver passage sequence)
    - ❑ Train end-to-end on the downstream tasks



Bridging the Gap: Results

Model Metric	NQ EM	HotpotQA EM	Email BLEU	Book BLEU
Naïve	33.07	28.01	5.57	11.5
Random	43.71	26.1	8.55	8.61
GTR	43.79	25.8	9.76	8.75
PSR	43.6	25.51	9.08	9.14
BGM	**45.37**	**35.64**	**10.42**	**12.07**

No external information

Randomized GTR retriever

GTR retriever

GTR + Reranker

SoTA < BGM

BGM is effective in adapting retrieved passages

Naïve is not always the worst

LLM already possesses a substantial amount of relevant knowledge (e.g., Book is from Amazon review)

GTR < BGM

HotpotQA is sensitive to irrelevant passages and has the most improvement

NQ typically only requires one retrieved passage, so the improvement is less

PSR < BGM

Pure reranking is not sufficient. Selection must also be taken into account.

(PSR) Atlas: Few-shot Learning with Retrieval Augmented Language Models, Izacard et al., arXiv 2022

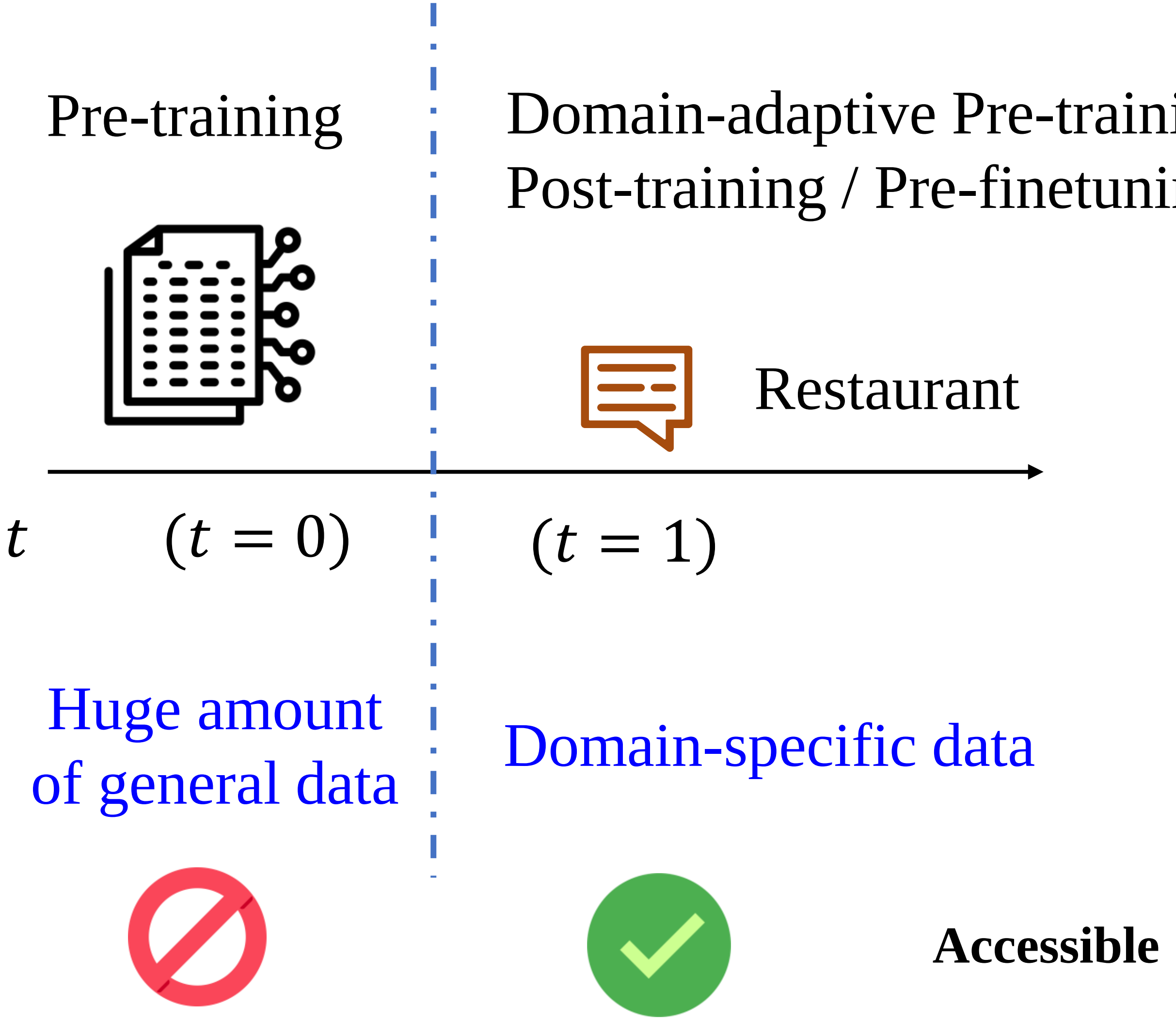
LLMs in A Dynamic World

How to adapt LLMs for the **dynamic world?**

Retrieval-augmented may not solve all problems (active research!). Another way is to **update** the parameters of LLMs with emerging data

This is, continual learning: (1) mitigate **forgetting**; and (2) encourage **knowledge transfer**

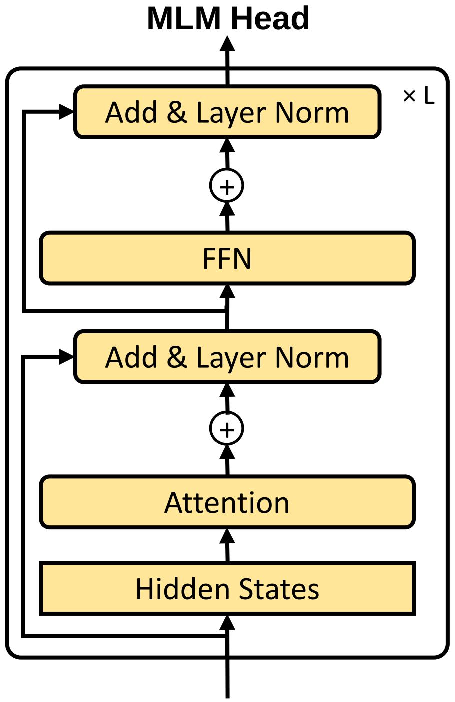
Post-training of Language Models



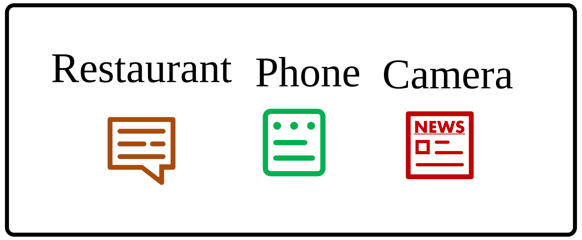
Two Needs:

- ❑ Due to **polysemy**, LM should be specialized or adapted to the target domain (**existing methods' focus**, may destroy useful general knowledge)
- ❑ General pre-trained knowledge should be preserved (**our focus**, a more informed adaptation that identifies what should be **preserved** and what should be **updated**)

Adapting a Language Model While Preserving its General Knowledge, Ke et al., EMNLP 2022

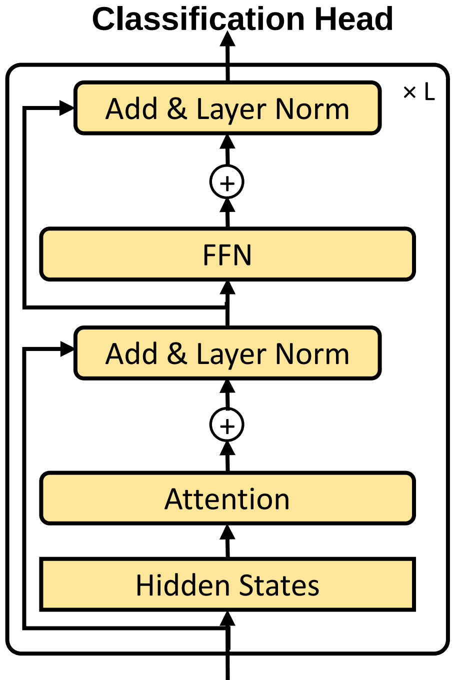


(A) Post-training



First, we post-train on **a specific domain**

(We use RoBERTa in this work)



(B) Fine-tuning

After (A), the performance is **evaluated** by end-tasks

End-tasks

- ASC-Restaurant
- ASC-Phone
- ASC-Camera

Each end-task **corresponding** to one domain and has its **own** training and testing set.

ASC: Aspect Sentiment Classification

Post-training of Language Model

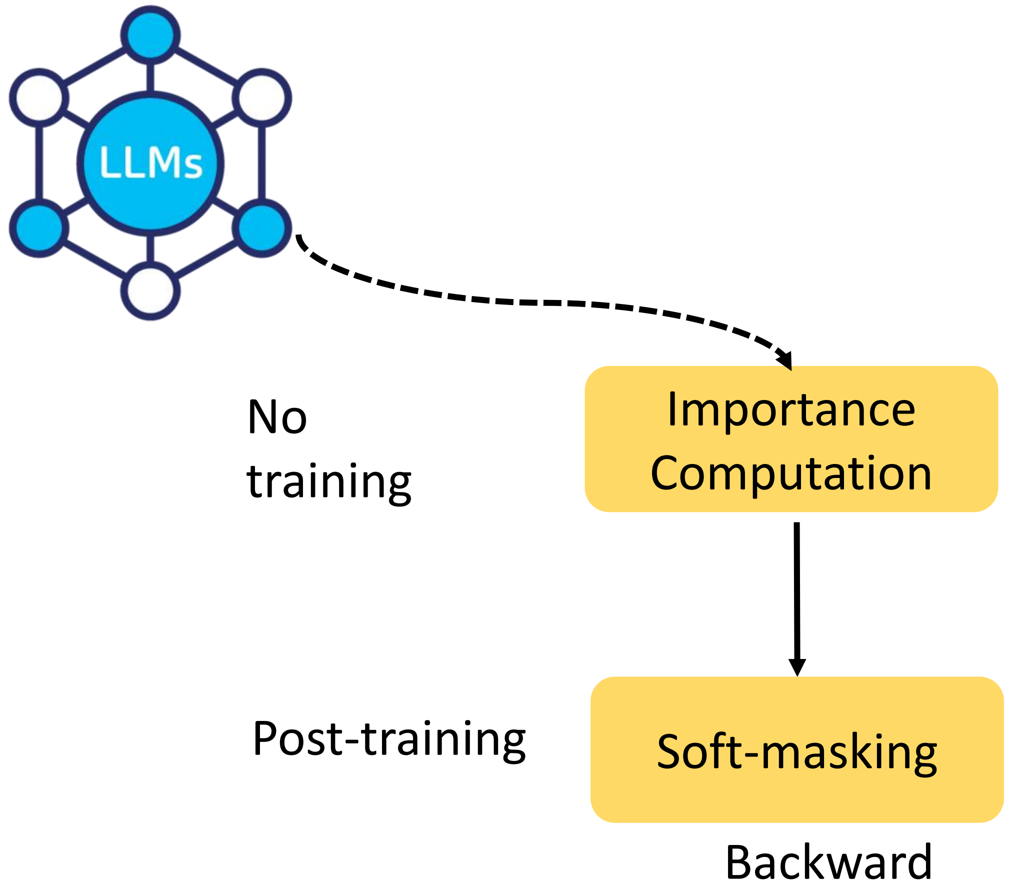
Unlabelde Domain Datasets			End-Task Classification Datasets				
Source	Dataset/Domain	Size	Dataset/Domain	Task	#Training	#Testing	#Classes
Reviews	Yelp Restaurant	758MB	Restaurant	Aspect Sentiment Classification (ASC)	3,452	1,120	3
	Amazon Phone	724MB	Phone	Aspect Sentiment Classification (ASC)	239	553	2
	Amazon Camera	319MB	Camera	Aspect Sentiment Classification (ASC)	230	626	2
Academic Papers	ACL Papers	867MB	ACL	Citation Intent Classification	1,520	421	6
	AI Papers	507MB	AI	Relation Classification	2,260	2,388	7
	PubMed Papers	989MB	PubMed	Chemical-protein Interaction Prediction	2,667	7,398	13

6 domains

post-training

Fine-tuning

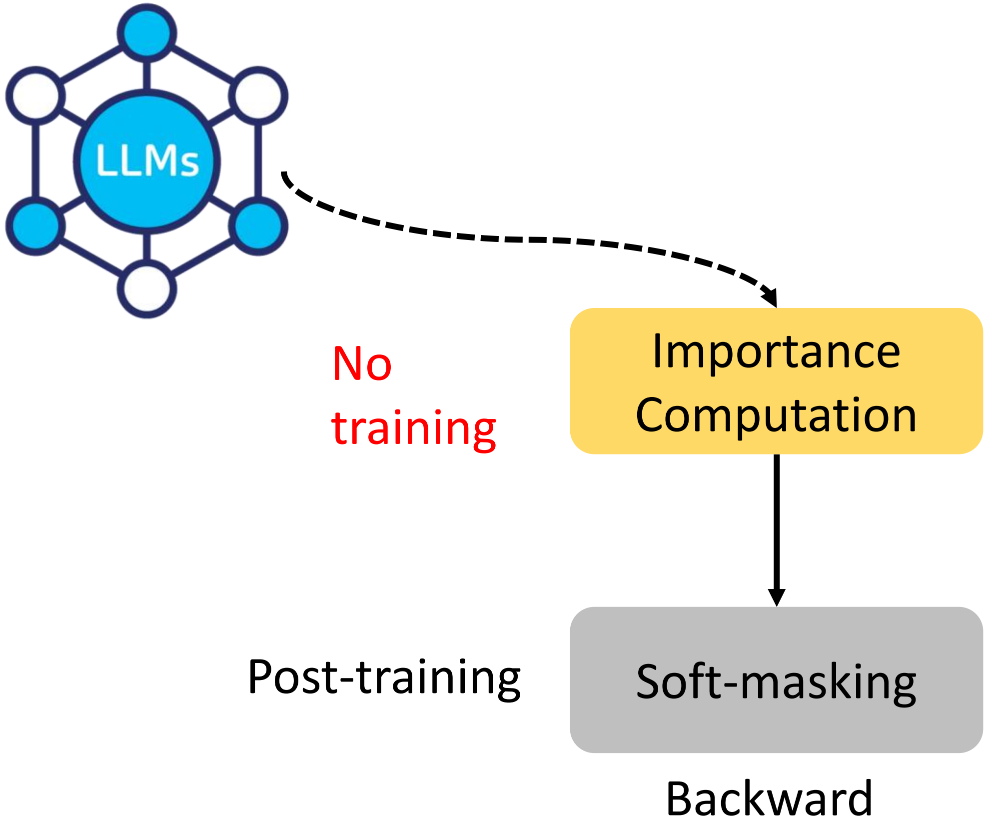
Post-training of Language Models



- ❑ **Detect** importance of units for general knowledge
- ❑ **Soft-masking** the important units in post-training

- ❑ **How to detect** importance of general knowledge
- ❑ **How to convert** the importance into soft-masks

Post-training of Language Models

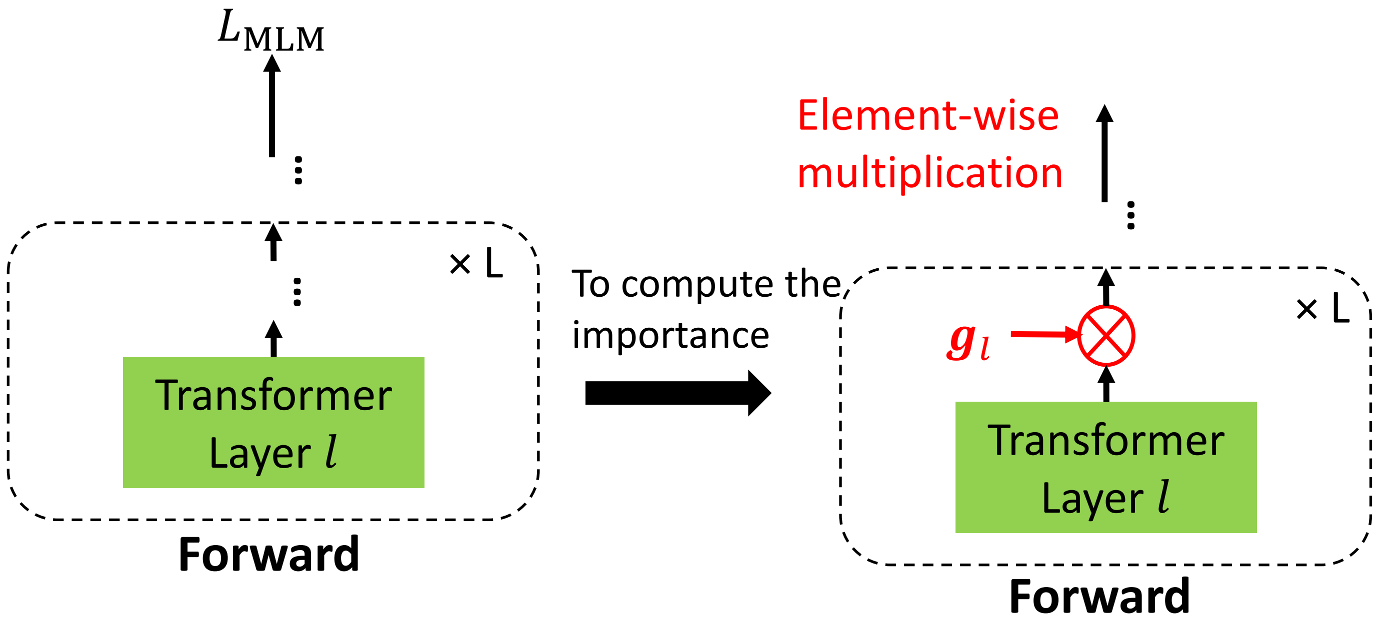


Goal: Compute the importance of units for **general** knowledge

Why?
1) Not all units are important
2) Given the important units, we can protect them afterward

No training involved. We only need the importance

Importance Computation



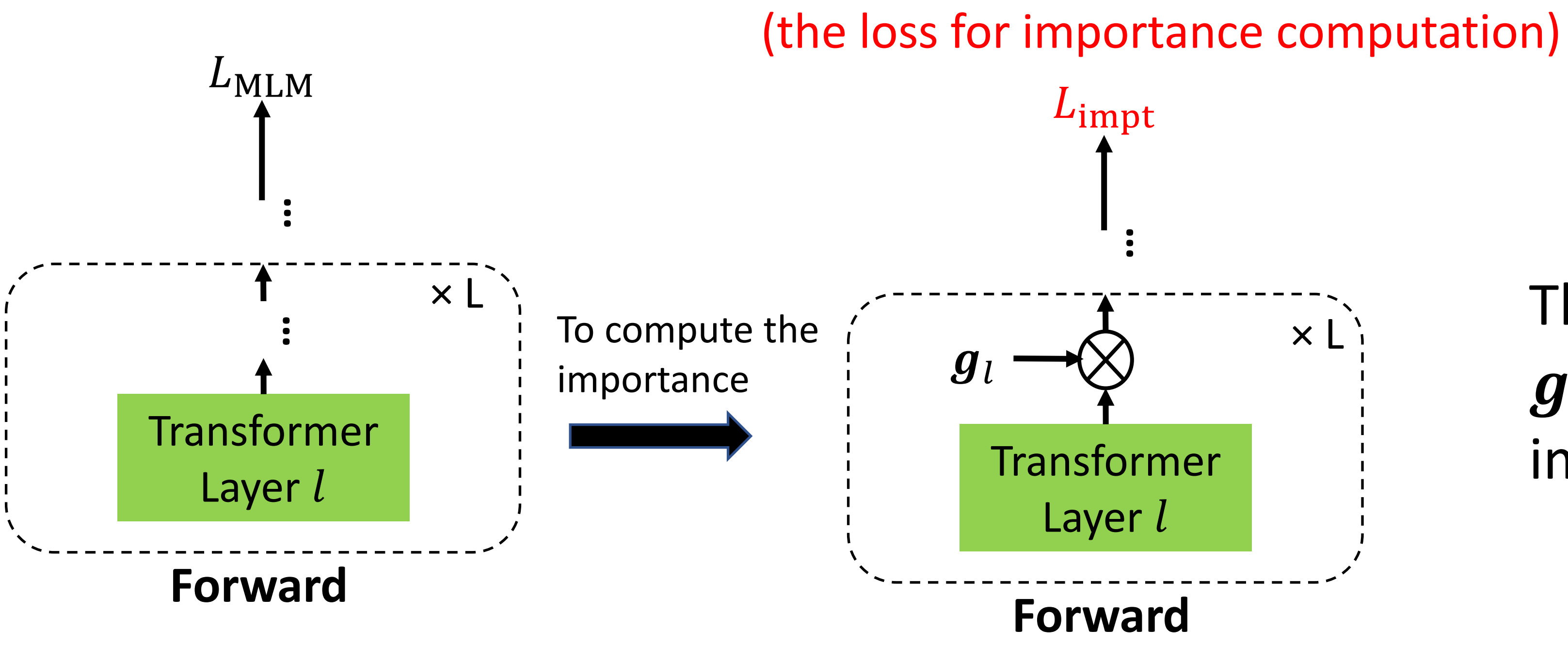
First, we added **virtual parameters** $\boldsymbol{g}_l$.

Each virtual parameter $g_{l,i}$ in $\boldsymbol{g}_l$ corresponding to an attention head or neurons (units)

It is **initialized as all 1's** and has its gradient but will **never change**.

Why? We only use its gradient to compute importance

Importance Computation



The gradient of $\boldsymbol{L}_{impt}$ w.r.t $\boldsymbol{g}_l$ will be used to compute importance.

Importance Computation

Due to **randomness**, same input will result in different output representation
Their distance indicates the **robustness**

$\boldsymbol{x}_m$ $\boldsymbol{x}_m$

Units that are important to the robustness

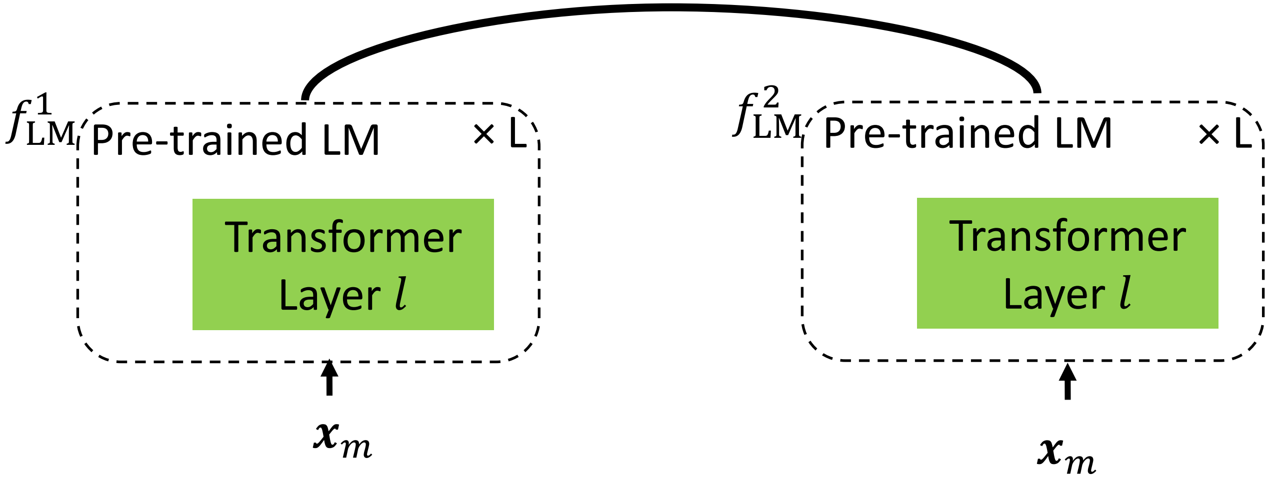
their changes will cause the pre-trained LM to change significantly

Units that are important to the pre-trained/general knowledge

So, the distance can be used as a **proxy** for general knowledge!

Importance Computation

$$L_{\text{impt}} = \text{KL}(f^1_{\text{LM}}(\boldsymbol{x}_m), f^2_{\text{LM}}(\boldsymbol{x}_m))$$



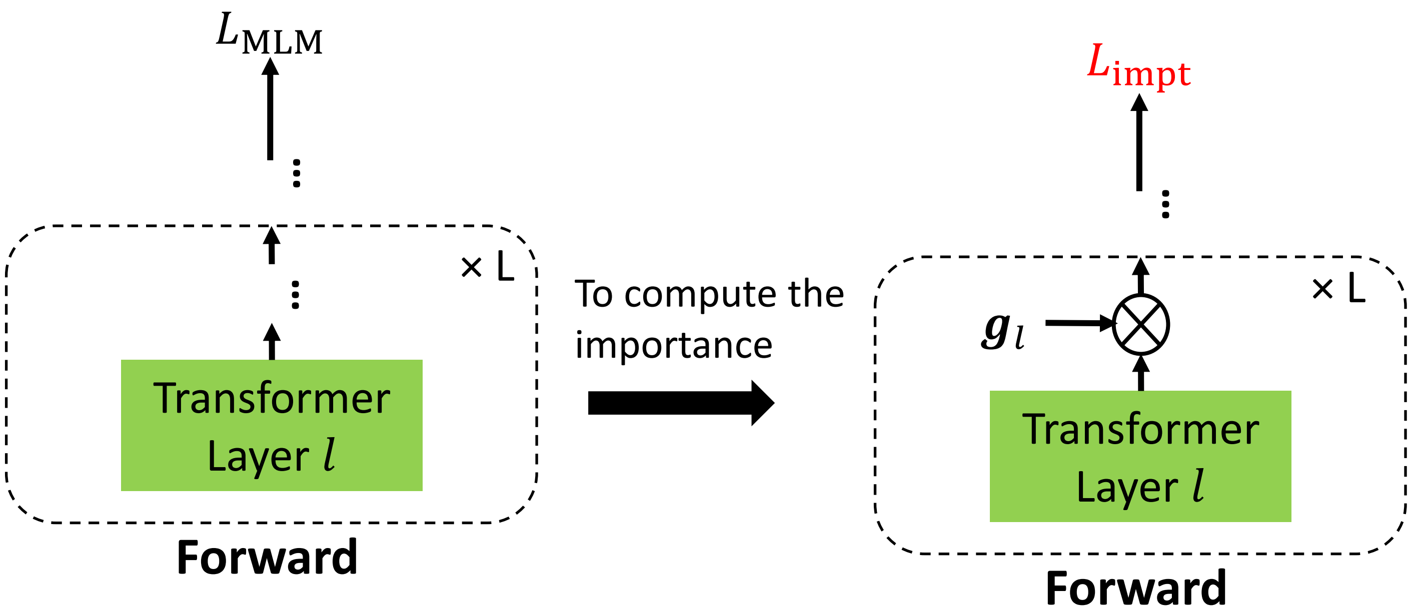
Based on the intuition, we propose another L_{impt}, which does not need pre-training data

KL: how different given two representations

$\boldsymbol{f^1_{LM}}$/ $\boldsymbol{f^2_{LM}}$: Transformer with different dropouts

$\boldsymbol{x_m}$: The domain data

Importance Computation



For **general knowledge,**

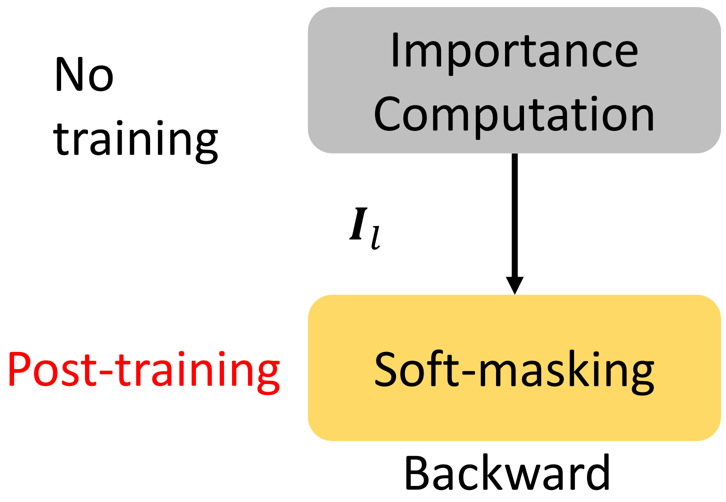
$$L_{\text{impt}} = \text{KL}(f_{\text{LM}}^{1}(\boldsymbol{x}_m), f_{\text{LM}}^{2}(\boldsymbol{x}_m))$$

$$\boldsymbol{\nabla}_{\boldsymbol{g}_l}^{m} = \frac{\partial L_{\text{impt}}(\boldsymbol{x}_m)}{\partial_{\boldsymbol{g}_l}}$$

$$\boldsymbol{I}_l = \frac{1}{M} \sum_M |\boldsymbol{\nabla}_{\boldsymbol{g}_l}^{m}|$$

Importance of units for general knowledge

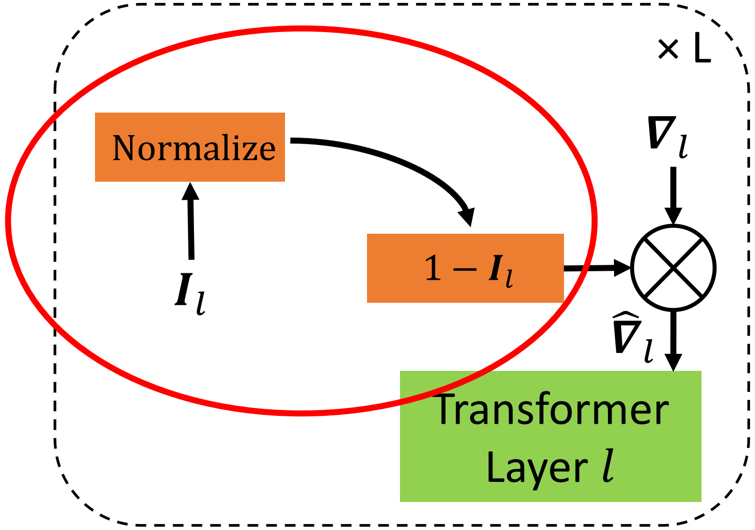
Post-training of Language Models



Goal: Soft-mask the **gradient** based on the importance

Why?
1) We need to protect them when training a new domain
2) We want to encourage knowledge transfer

Soft-masking



Backward

First, we normalize the importance so that they are comparable

$$\boldsymbol{I}_l = |\mathrm{Tanh}(\mathrm{Norm}(\boldsymbol{I}_l))|$$

make sure the importance is [0,1]

Next, we soft-mask the gradient (in backward pass)

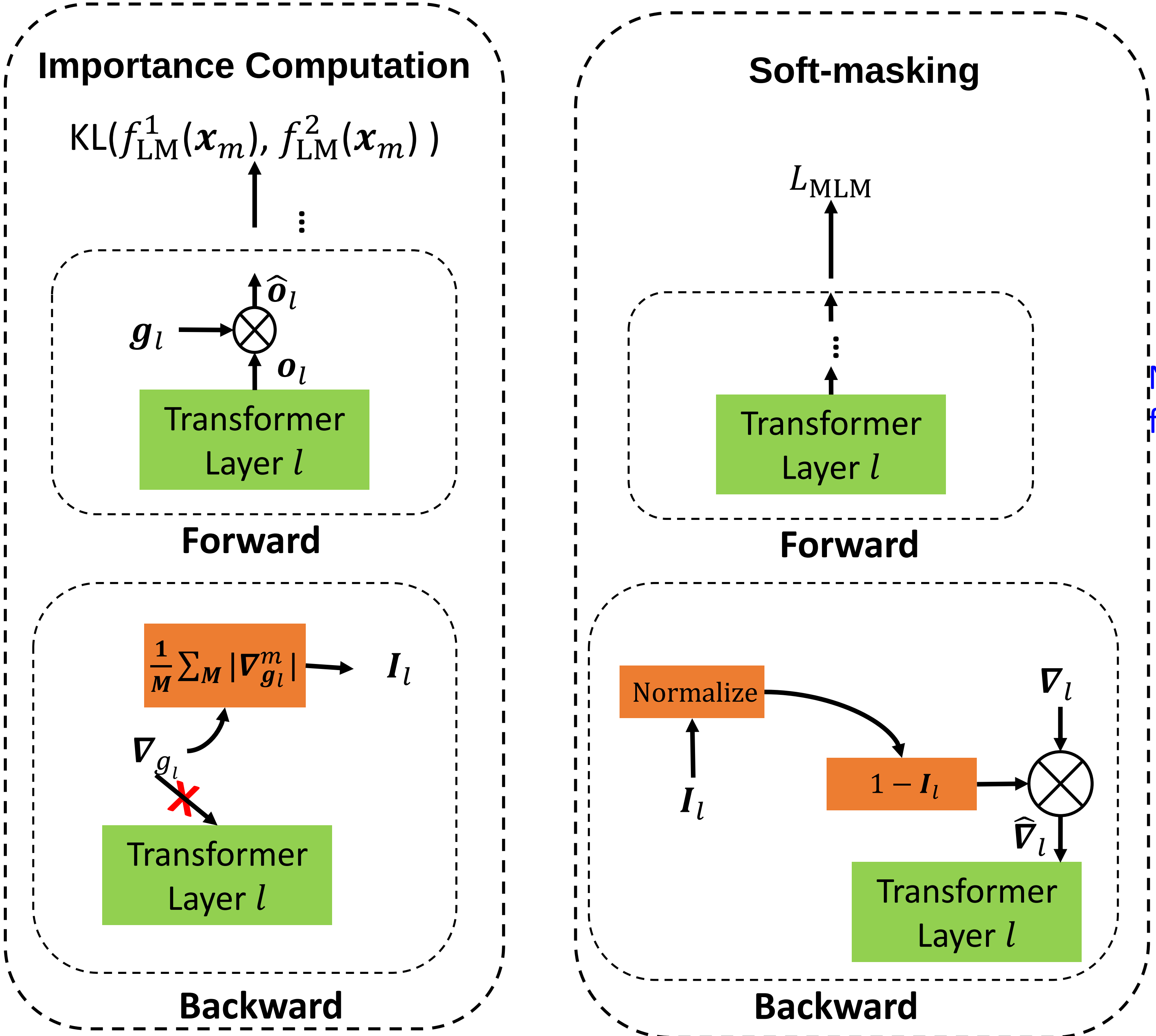
$$\boldsymbol{\nabla}'_l = (1 - \boldsymbol{I}_l) \otimes \boldsymbol{\nabla}_l$$

This only affects the backward pass
so forward KT and full LM are still possible
Not only provides protection, but also allow knowledge transfer

KL loss as L_{impt}

Use gradient to indicate importance, but the gradient does not optimize the layer

I_l indicates the importance for general knowledge



Nothing changed in forward pass

Use the importance to soft-mask the backward gradient flow

		Camera	Phone	Restaurant	AI	ACL	PubMed	Average
	No post-train	78.82	83.75	79.81	60.98	66.11	72.38	73.64
	MLM	84.39	82.59	80.84	68.97	68.75	72.84	76.4
	MLM (Adapter)	83.62	82.71	80.19	60.55	68.87	71.68	74.6
SoTA post-training baselines	MLM+KD	82.79	80.08	80.4	67.76	68.19	72.35	75.26
SoTA post-training baselines	MLM+AdaptedDeiT	86.86	83.08	79.7	69.72	69.11	72.69	76.86
SoTA post-training baselines	MLM+SimCSE	84.91	83.46	80.88	69.1	69.89	72.77	76.84
SoTA post-training baselines	MLM+TaCL	81.98	81.87	81.12	64.04	63.18	69.46	73.61
	DGA	**88.52**	**85.47**	**81.83**	**71.99**	**71.01**	**73.65**	**78.74**

- w/o Pre-trained < MLM
  Not surprising, as post-training has been demonstrated to improve performance in the literature.

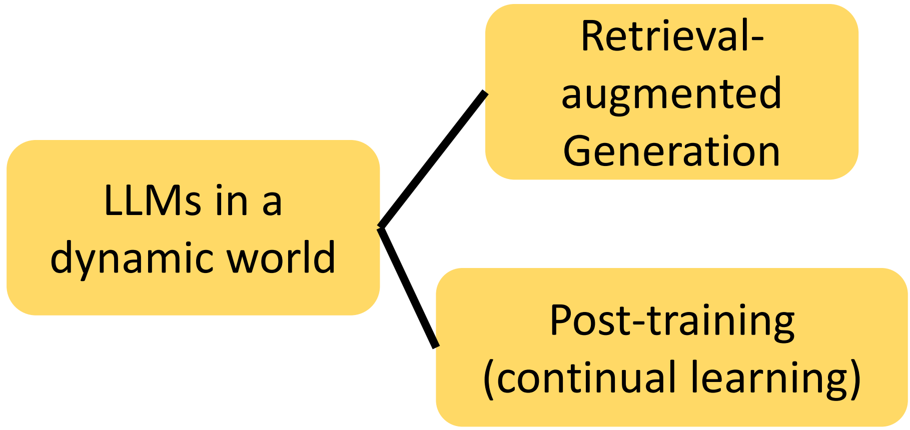
- MLM (Adapter) < MLM
  Efficient tuning like adapter may not have sufficient trainable parameters for post-training

- w/o Pre-trained < MLM < SoTA < DGA
  DGA is better than pure MLM and SoTA post-training. DGA can not only mitigate forgetting of the general knowledge but also adapt to suite the target domain

- SoTA < DGA
  SoTAs either only focus on preserving knowledge (KD), or adapting to the target domain, which are not enough

(Extended to continual pre-training on a sequence of domains)
Continual Pre-training of Language Models, Ke et al., ICLR 2023

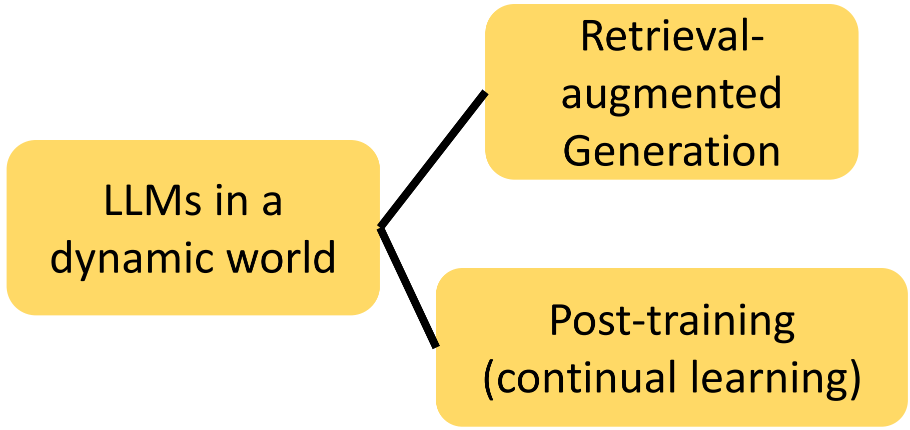
Adapting LLMs for A Dynamic World



A more **ambitious** vision is to make LLMs **fully autonomous**, which requires LLMs to **self-initiate** and **adapt to new circumstances,** so that the AI system can **independently** acquire new knowledge.

My vision: humans **are intrinsically motivated by novelty** to learn; same principle can also apply to AI system!

An Example of Autonomy



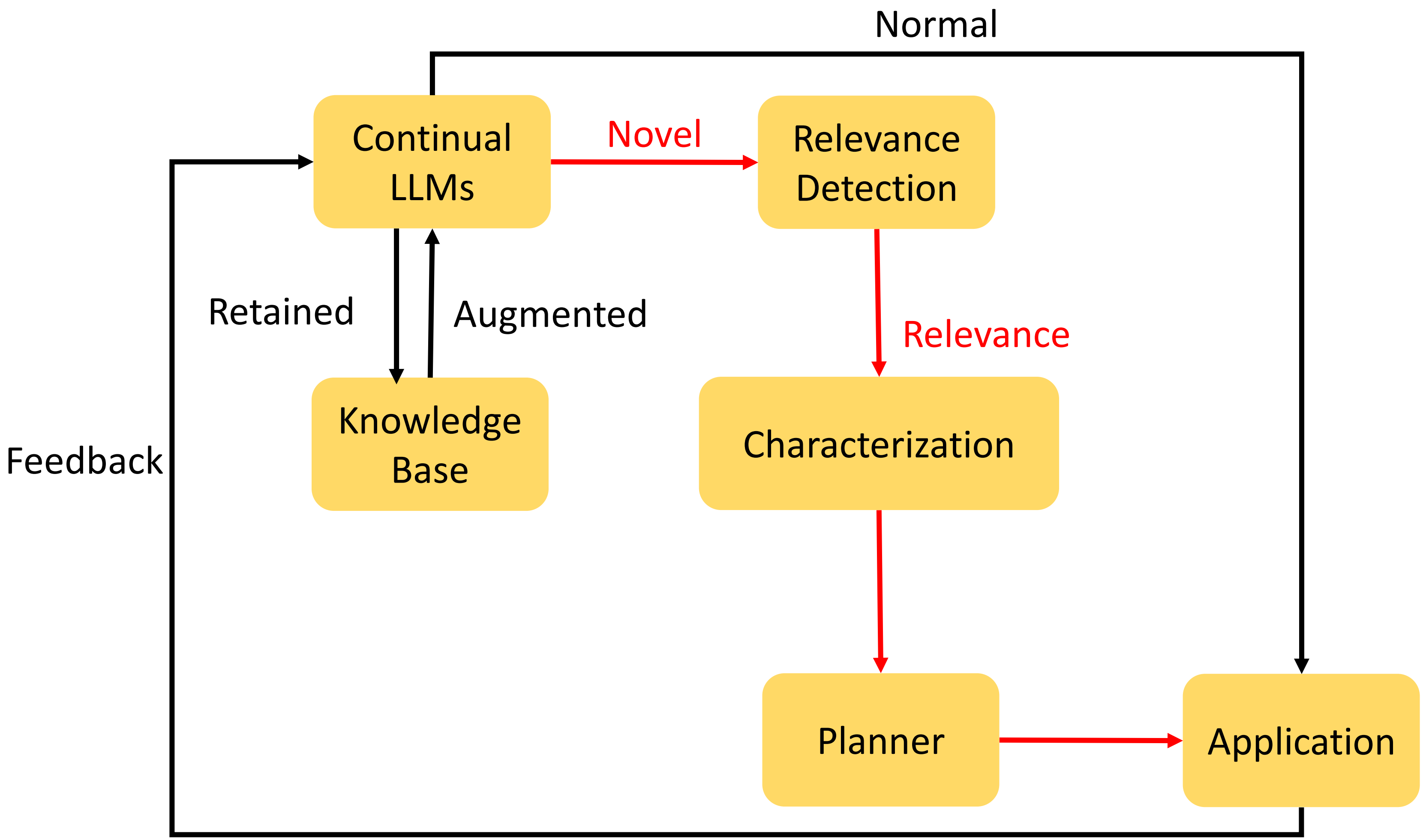
User: Finish the sentence in Vincent's tone
System: Sorry, I didn't fully understand, do you mean:

Option-1: Vincent as the artist Vincent Van Gogh?
Option-2: Any specific person called Vincent? It would be good if you could provide more information

In this example, the system

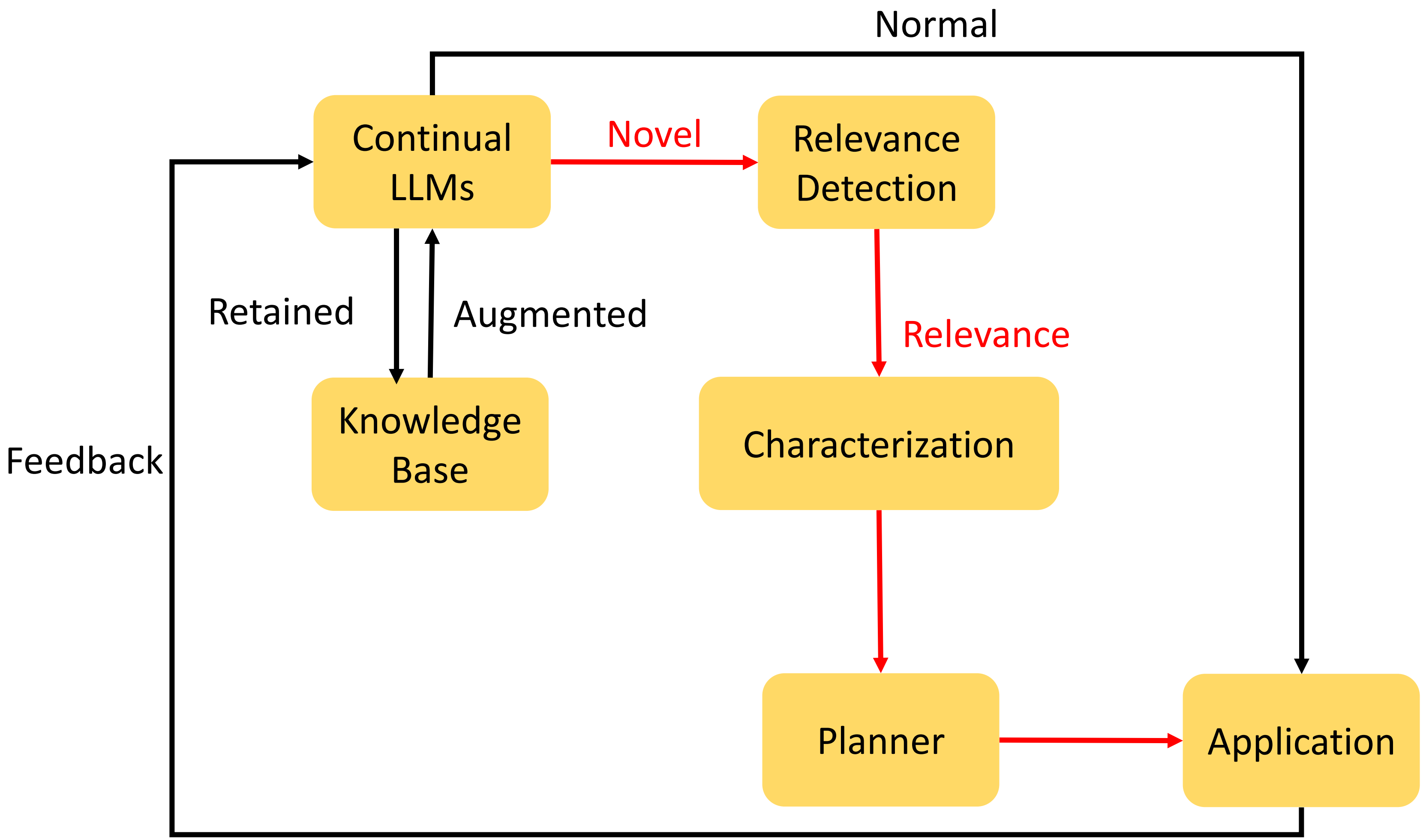
- **Encounters** a novel prompt (i.e., novelty) that the agent does not understand or there is ambiguity
- **Identifies** which aspects it understands, or which aspect is challenging (i.e., characterization)
- **Adapts** by posing questions or offering choices (i.e., adaptation)

A Possible Framework



- **Continual LLMs** to detect novelty (if the input is normal, it can simply give output to the application)
- **Relevance detection** to check whether the novelty is relevant to the task it is focused on
- **Characterization** to identifying understandable and unclear parts
- **Planner** to generate a strategy for responses, e.g., asking questions to user
- **Feedback** needs to be continually integrated
- **Knowledge base** may be needed to augment and retain essential knowledge

A Possible Framework



Most existing works are dedicated to the black part, which includes active research areas like **retrieval-augmented generation** and **continual learning**.

The other components remain largely unexplored!

KEY TAKE AWAY

Adapting LLMs for A Dynamic World

Active Research

Retrieval-augmented Generation

Post-training (continual learning)

Ambitious goal

Fully autonomous LLMs


THANK
YOU!