Continual Learning of Language Models

Haowei Lin

Peking University

2023/08/08

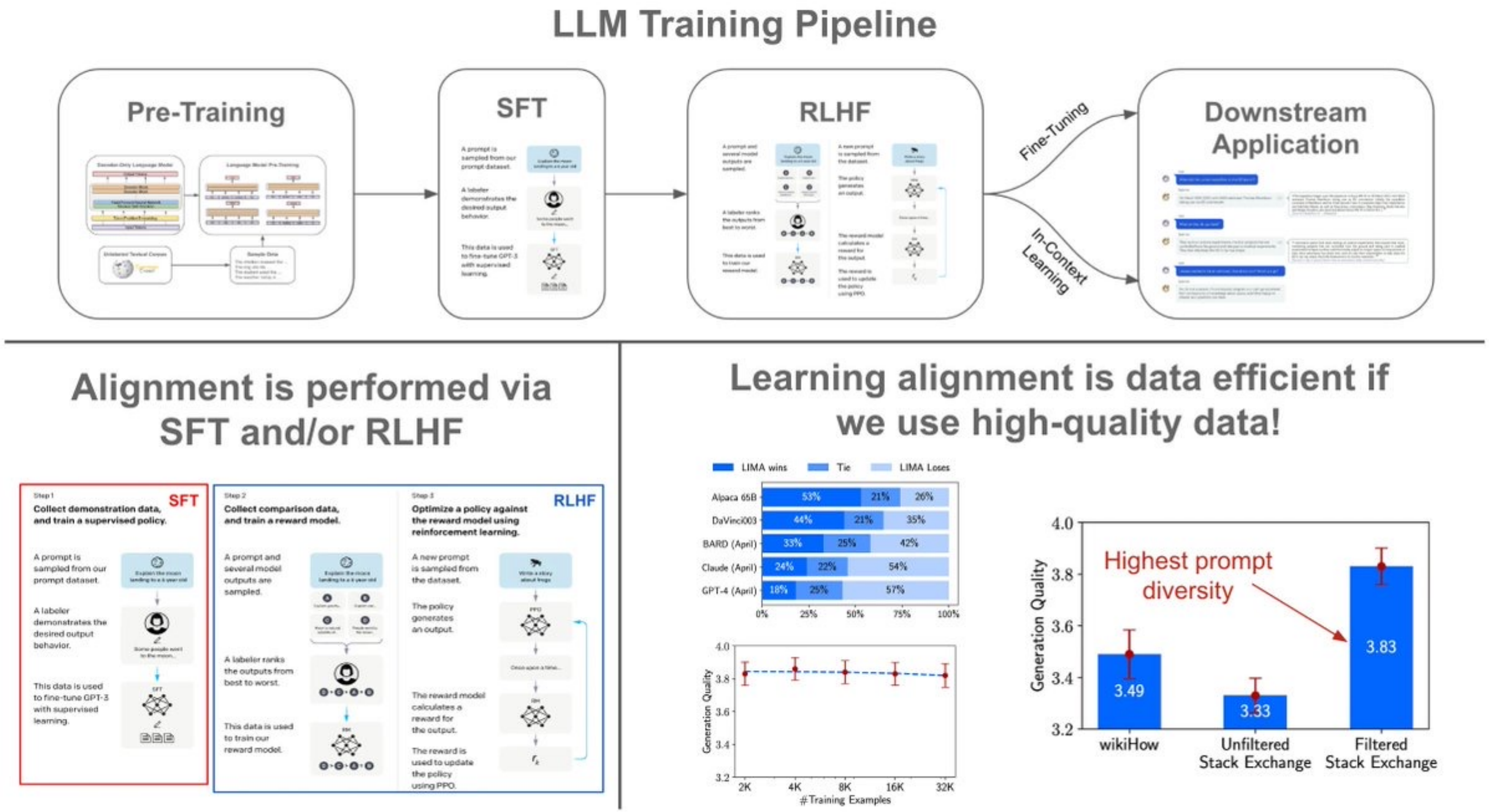
Overview

- From RLHF to Continual Learning
- A quick Introduction to Traditional Continual Learning
- Continual Learning of LMs
- Open Questions

Overview

- From RLHF to Continual Learning
- A quick Introduction to Traditional Continual Learning
- Continual Learning of LMs
- Open Questions

The consensus is that …



Pretraining Language Models with Human Preferences

- PHF: pretraining with human feedback

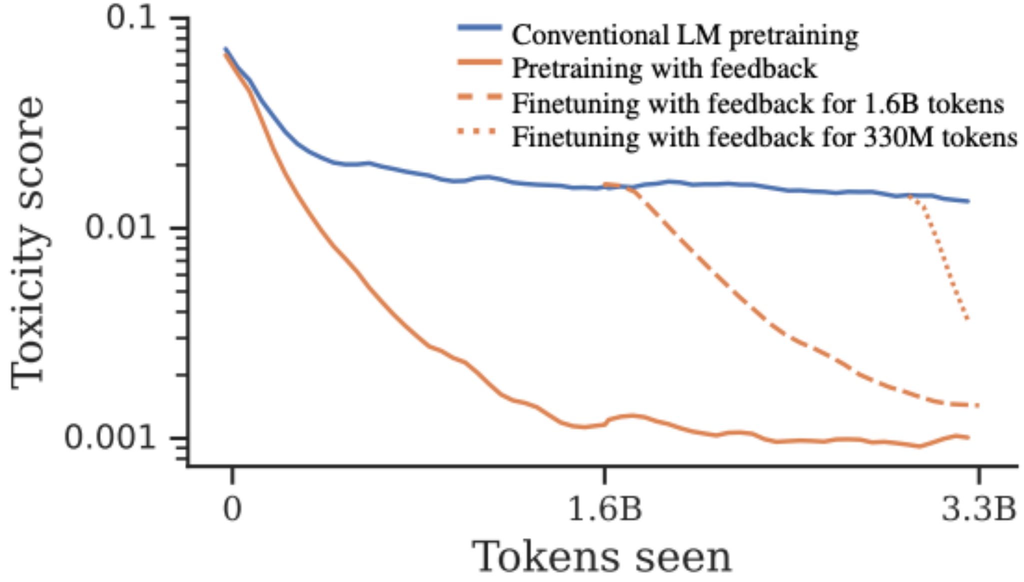


Figure 1: Toxicity score (lower is better) of LMs pretrained with the standard objective (solid blue), using conditional training (solid orange) and LMs finetuned using conditional training for 1.6B (orange dashed) and 330M tokens (orange dotted). Pretraining with Human Feedback (PHF) reduces the amount of offensive content much more effectively than finetuning with human feedback.

Alignment should happen at the pre-training phase…

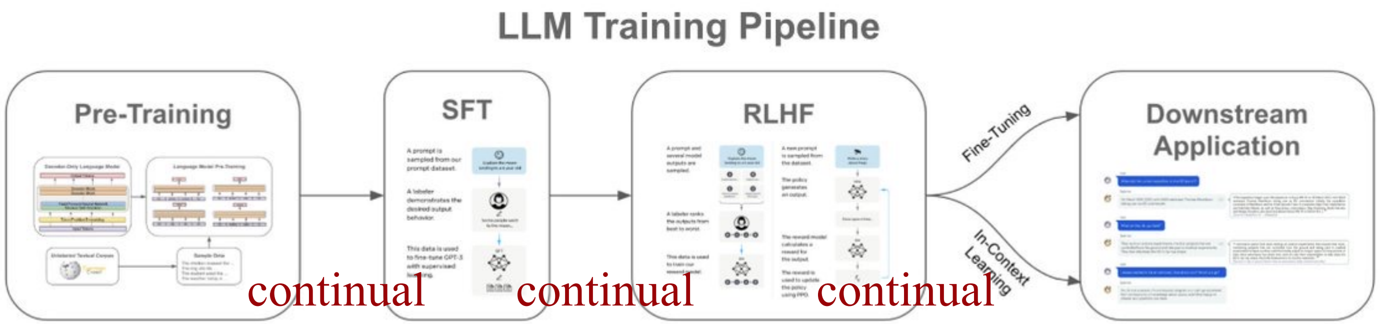
But it seems impossible for us to pre-train LLMs from scratch…?

Korbak, Tomasz, et al. "Pretraining language models with human preferences." *International Conference on Machine Learning*. PMLR, 2023.

“Alignment”

When doing alignment, you also care about forgetting (helpfulness)

- More general: adaptation
  - There is already a system, and we want it to be capable of new tasks
  - Alignment: adapt AI systems / LLMs to become human-friendly systems
- Other related topics:
  - Transfer learning, domain adaptation
- The adaptation happens many times!
  - Continual learning



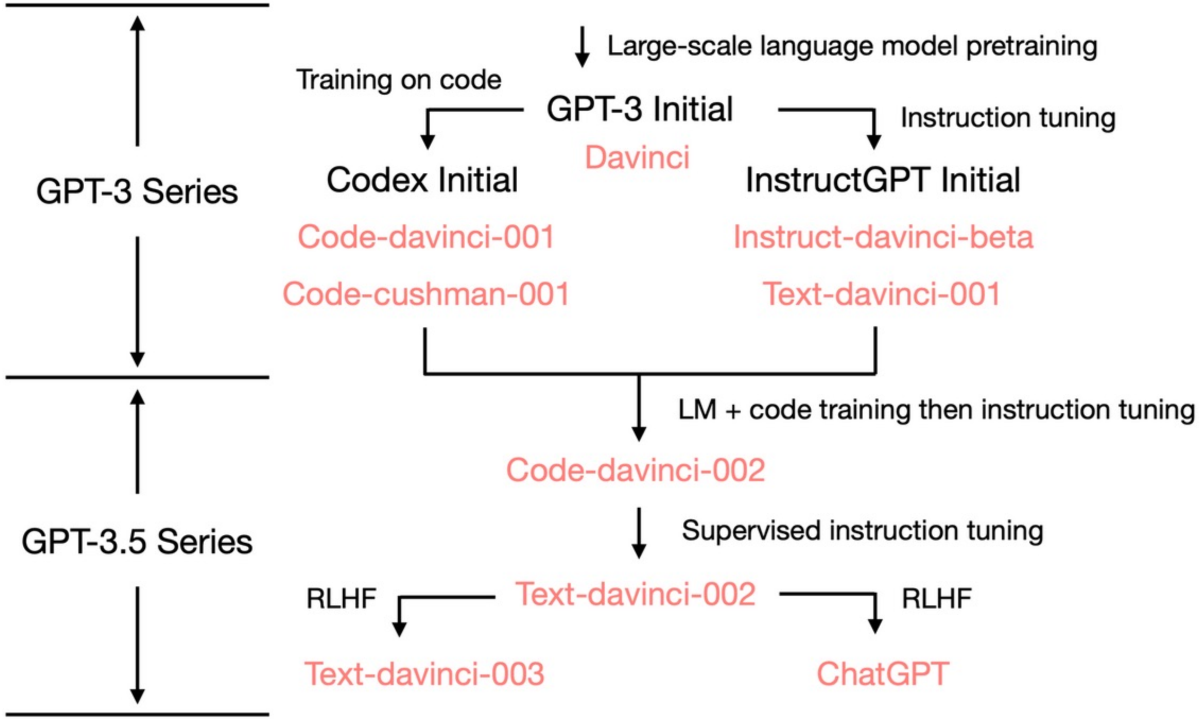
continual learning

continual learning

continual learning

Human and LLMs are continual learners

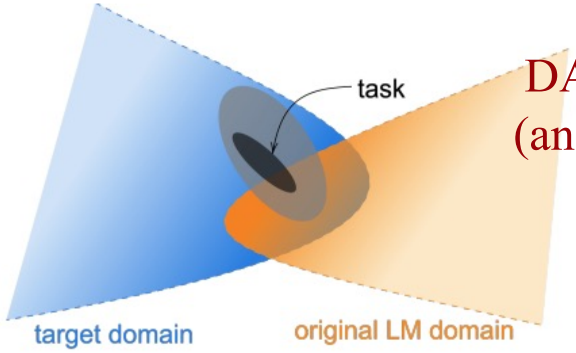
- Model patching & continual training of LLMs are important
  - That's how OpenAI successfully trained GPTs (version control, clever incremental updating, maintenance)



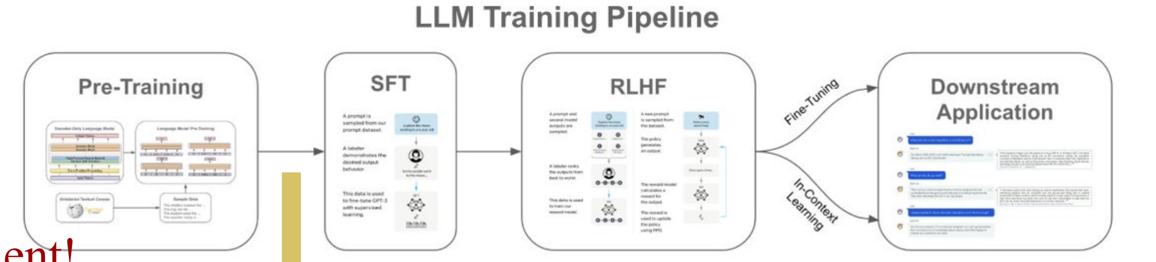
Yao Fu "How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Sources"

Another consensus (but ancient): DAPT

- Domain Adaptive Pre-training



DAPT will bring improvement!
(another continual learning case)

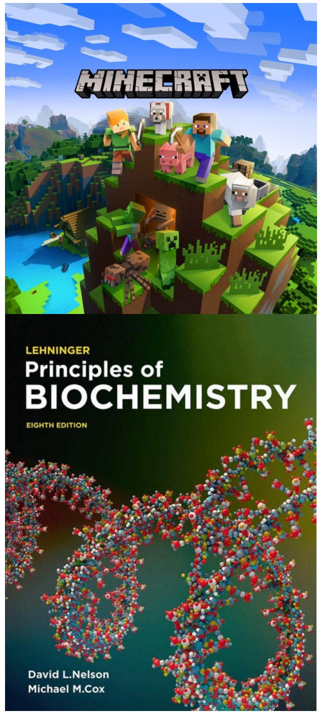


DAPT

Domain	Pretraining Corpus	# Tokens	Size	$\mathcal{L}_{\text{ROB.}}$	$\mathcal{L}_{\text{DAPT}}$
BIOMED	2.68M full-text papers from S2ORC (Lo et al., 2020)	7.55B	47GB	1.32	0.99
CS	2.22M full-text papers from S2ORC (Lo et al., 2020)	8.10B	48GB	1.63	1.34
NEWS	11.90M articles from REALNEWS (Zellers et al., 2019)	6.66B	39GB	1.08	1.16
REVIEWS	24.75M AMAZON reviews (He and McAuley, 2016)	2.11B	11GB	2.10	1.93
ROBERTA (baseline)	see Appendix §A.1	N/A	160GB	‡1.19	-

Dom.	Task	ROBA.	DAPT	¬DAPT
BM	CHEMPROT	81.9$_{1.0}$	**84.2**$_{0.2}$	79.4$_{1.3}$
	†RCT	87.2$_{0.1}$	**87.6**$_{0.1}$	86.9$_{0.1}$
CS	ACL-ARC	63.0$_{5.8}$	**75.4**$_{2.5}$	66.4$_{4.1}$
	SCIERC	77.3$_{1.9}$	**80.8**$_{1.5}$	79.2$_{0.9}$
NEWS	HYP.	86.6$_{0.9}$	**88.2**$_{5.9}$	76.4$_{4.9}$
	†AGNEWS	**93.9**$_{0.2}$	**93.9**$_{0.2}$	93.5$_{0.2}$
REV.	†HELPFUL.	65.1$_{3.4}$	**66.5**$_{1.4}$	65.1$_{2.8}$
	†IMDB	95.0$_{0.2}$	**95.4**$_{0.2}$	94.1$_{0.4}$

Domain	Task	Label Type	Train (Lab.)	Train (Unl.)	Dev.	Test	Classes
BIOMED	CHEMPROT	relation classification	4169	-	2427	3469	13
	†RCT	abstract sent. roles	18040	-	30212	30135	5
CS	ACL-ARC	citation intent	1688	-	114	139	6
	SCIERC	relation classification	3219	-	455	974	7
NEWS	HYPERPARTISAN	partisanship	515	5000	65	65	2
	†AGNEWS	topic	115000	-	5000	7600	4
REVIEWS	†HELPFULNESS	review helpfulness	115251	-	5000	25000	2
	†IMDB	review sentiment	20000	50000	5000	25000	2



Gururangan, Suchin, et al. "Don't stop pretraining: Adapt language models to domains and tasks." *ACL* (2020).

Overview

- From RLHF to Continual Learning
- A quick Introduction to Traditional Continual Learning
- Continual Learning of LMs
- Open Questions

Continual learning (CL)

- Learn from streaming experiences (may forget past knowledge)
  - **CL vs. Online learning**
    - No distributional shift in online learning
  - **CL vs. transfer learning**
    - Not continuous, the src is similar to tgt, only one directional: src helps tgt
    - E.g., ELMo, BERT, RoBERTa
  - **CL vs. multitask learning (MTL)**
    - MTL retains no knowledge except data
    - MTL is hard to relearn all task whenever a new task appears (you need to re-train models)
    - MTL is often considered as the upper bound of CL
    - E.g., Machine Translation
  - Nicknames: lifelong learning, incremental learning, never-ending learning

A quick Introduction to Traditional Continual Learning

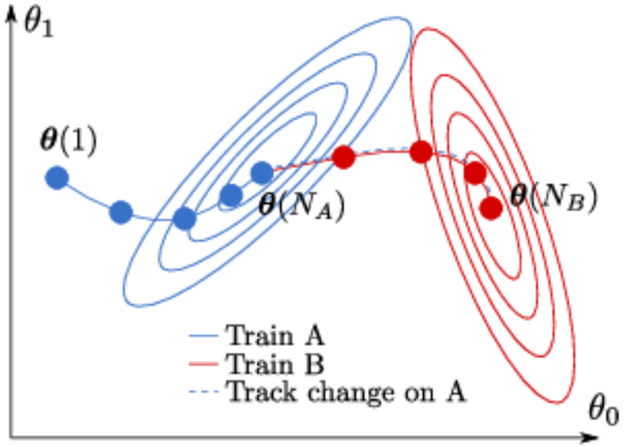
- Desiderata
- Settings
- Challenges
- Methods
- Applications

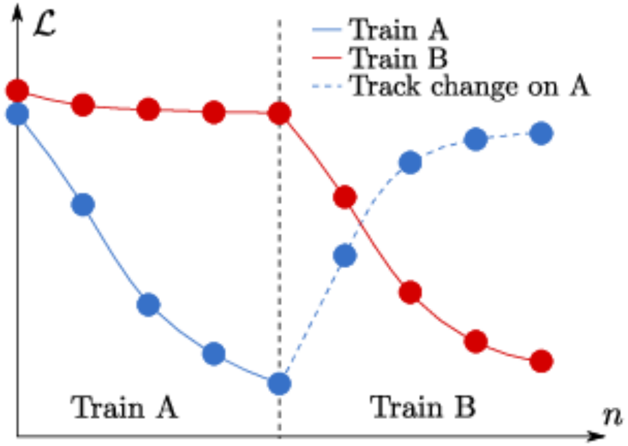
A quick Introduction to Traditional Continual Learning

- Desiderata
- Settings
- Challenges
- Methods
- Applications

Catastrophic forgetting

- Prevent **catastrophic forgetting** (CF)
  - French, Robert M. "Catastrophic forgetting in connectionist networks." Trends in cognitive sciences 3.4 (1999): 128-135.
  - Kirkpatrick, James, et al. "Overcoming catastrophic forgetting in neural networks." Proceedings of the national academy of sciences 114.13 (2017): 3521-3526.





Knowledge transfer

- Achieve positive **forward KT** and **backward KT**
  - Forward KT: old knowledge helps new tasks
  - Backward KT: new knowledge helps old tasks

Testing task

Tasks trained so far	T_1	T_2	T_3	T_4	T_5	......
T_1	$R_{1,1}$					
T_2	$R_{2,1}$	$R_{2,2}$				
T_3	$R_{3,1}$	$R_{3,2}$	$R_{3,3}$			
T_4	$R_{4,1}$	$R_{4,2}$	$R_{4,3}$	$R_{4,4}$		
T_5	$R_{5,1}$	$R_{5,2}$	$R_{5,3}$	$R_{5,4}$	$R_{5,5}$	
⋮	⋮					
	$R_{t,1}$	......				

If goes down -> forgetting
If goes up -> positive transfer

- Forward Transfer (FWT): $\frac{1}{T-1}\sum_{i=1}^{T-1} R_{i,i} - R_i$
- Backward Transfer (BWT): $\frac{1}{T-1}\sum_{i=1}^{T-1} R_{t,i} - R_{i,i}$

Lopez-Paz and Ranzato, Gradient Episodic Memory for Continual Learning, NIPS 2017

A quick Introduction to Continual Learning

- Desiderata
- Settings
- Challenges
- Methods
- Applications

Task/Class/Domain incremental learning

- **TIL**: task-ID is available during training and testing
  - Gaokao: study Chinese – math – English – Physics ...
  - When testing, you are told what you are doing
  - Note: in CL, we only build one model (memory overhead, human-like)
- **CIL**: task-ID is not available during testing
  - You learned how to classify <u>cats</u> and <u>dogs</u>, one day you learn how to classify <u>pigs</u> and <u>dogs</u>, then you should be able to classify three of them (w/o seeing cats again).
- **DIL**: when the label space is unified (usually no task-ID in testing)
  - E.g. sentiment classification (positive, negative) on Yelp, IMDB, Reddit
  - E.g., Generative model (the task is aways generation)

A quick Introduction to Continual Learning

- Desiderata
- Settings
- Challenges
- Methods
- Applications

Challenges

- **Stability-plasticity**
  - Preserving the learned knowledge vs. learning from new experiences
- **Transfer-interference**
  - Knowledge transfer vs. knowledge interference
  - Increase parameter-sharing is a common way towards KT, but…
  - Transfer is not always positive! Avoid negative transfer…
- **Task separation**
  - Mostly in CIL and DIL, it's hard to predict task-ID
  - In learning the current experience, the learner cannot see previous or future data, thus it's hard to establish decision boundaries between tasks

A quick Introduction to Continual Learning

- Desiderata
- Settings
- Challenges
- Methods
- Applications

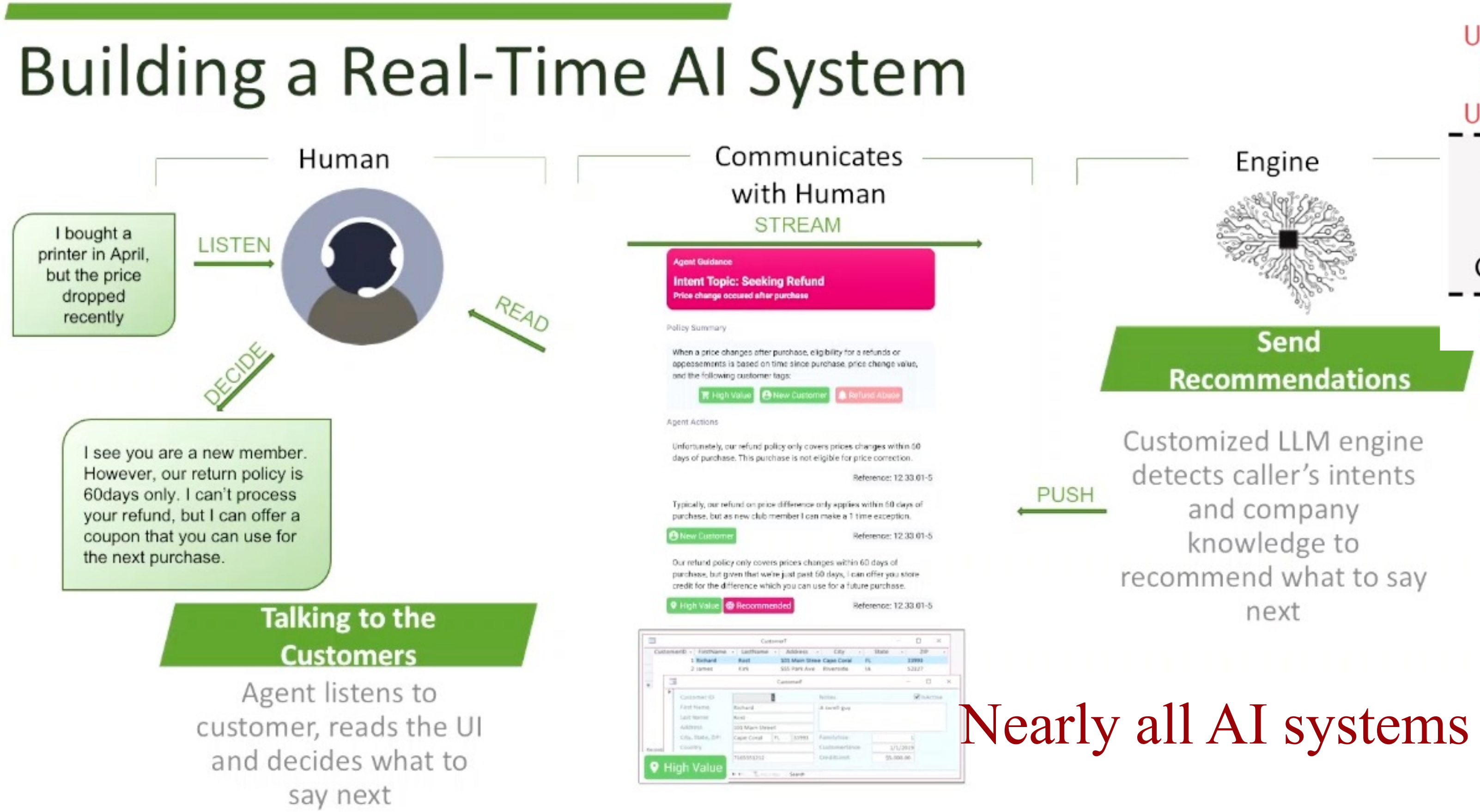
Methods: very very brief!

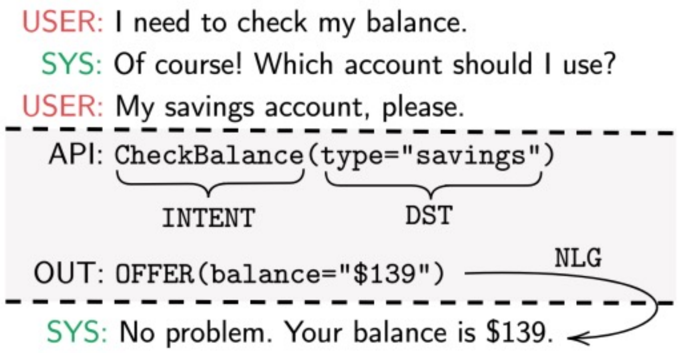
- \- CF, + KT
  - **Regularization-based**: regularize the model / feature / output space
    - E.g., using old model to distill new model, orthogonal projection of gradient
  - **Replay-based**: save (or generate) a small amount of past data
    - E.g., experience replay, pseudo replay
  - **Architecture-based**: build sub-networks inside the whole network
    - E.g., parameter isolation (no forgetting in TIL, no KT), modular network

A quick Introduction to Continual Learning

- Desiderata
- Settings
- Challenges
- Methods
- Applications

Applications: Task-oriented dialog system



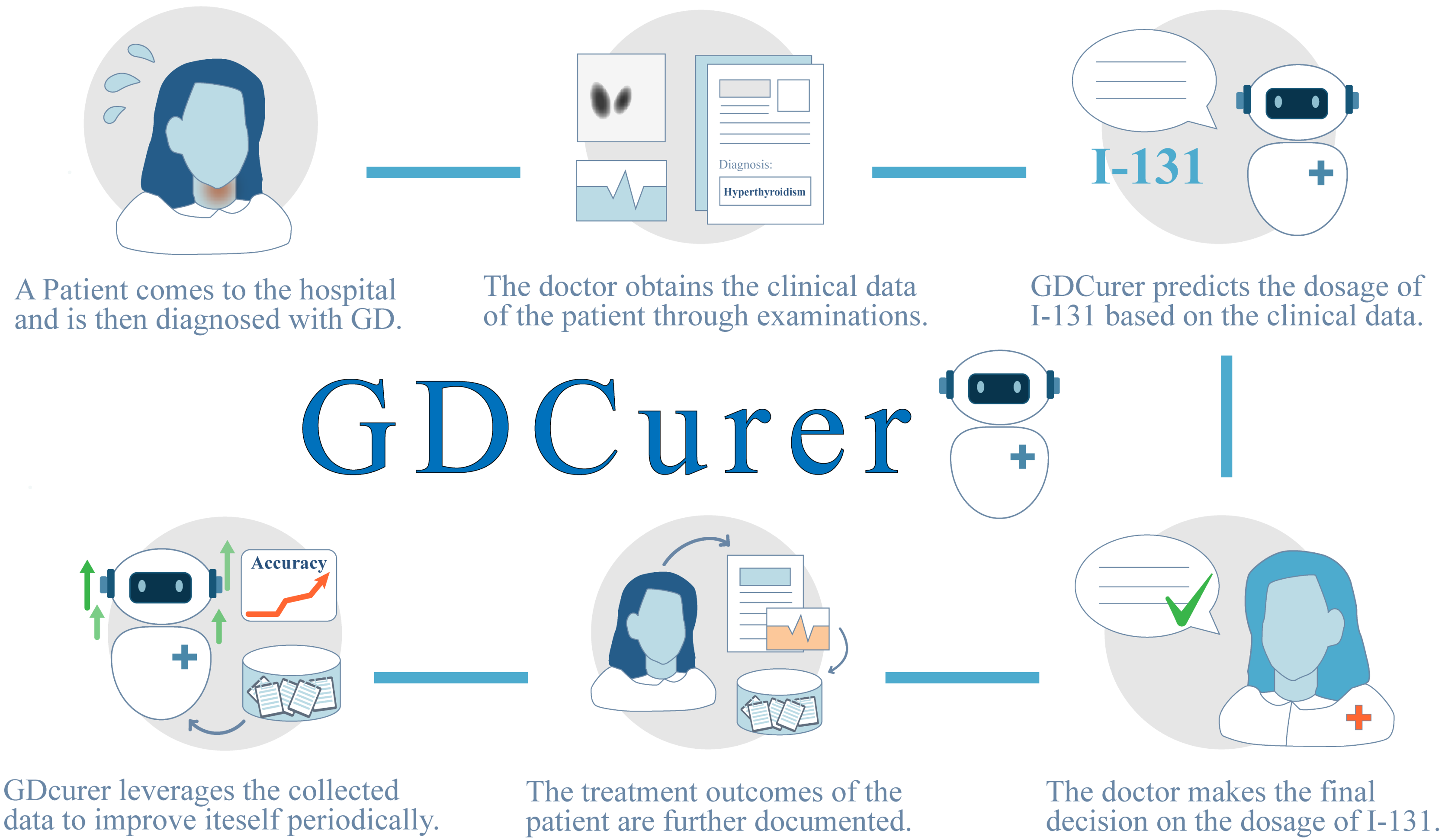


Nearly all AI systems needs continual learning!

BirchAI

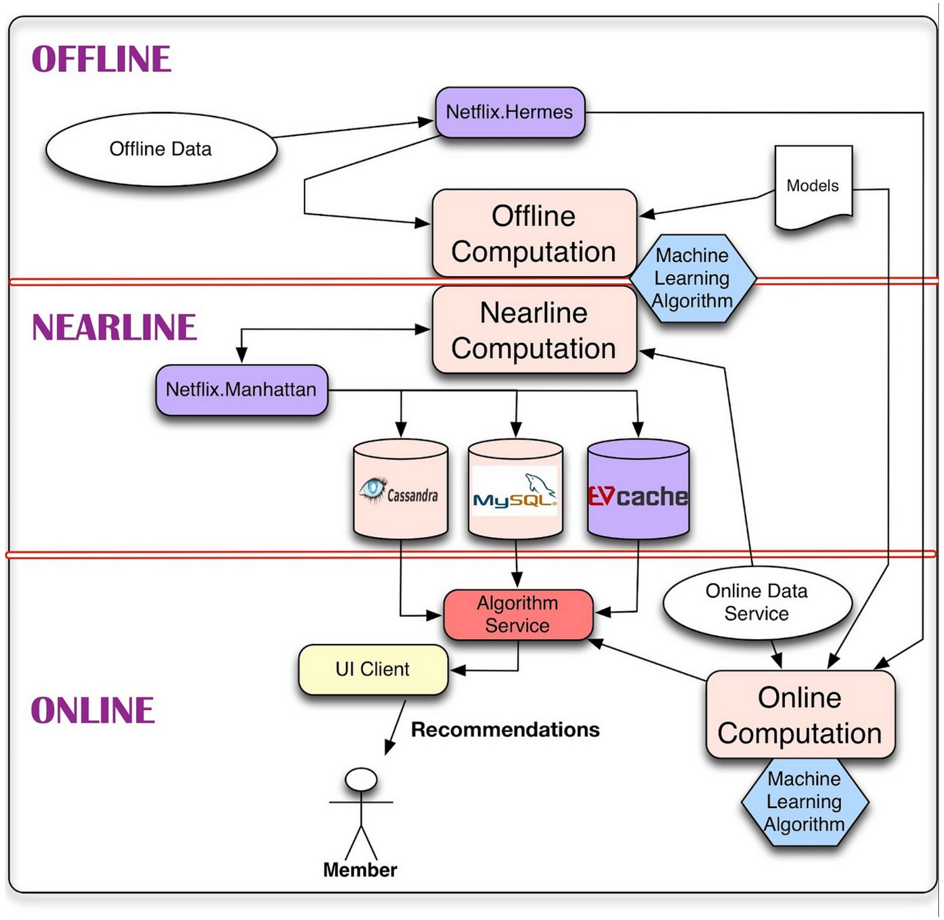
Andrea Madotto et al, Continual Learning in Task-Oriented Dialogue Systems, EMNLP (2021)
Yinhan Liu, Build an AI system: Applying Reinforcement learning with human feedback (RLHF) on LLM to advance customization

Applications: medical applications



Haowei Lin, et al. “GDCurer: An AI-assisted Drug Dosage Prediction System for Graves’ Hyperthyroidism”

Applications: Recommender systems



Search engines require CL, too.

[System Architectures for Personalization and Recommendation | by Netflix Technology Blog | Netflix TechBlog]

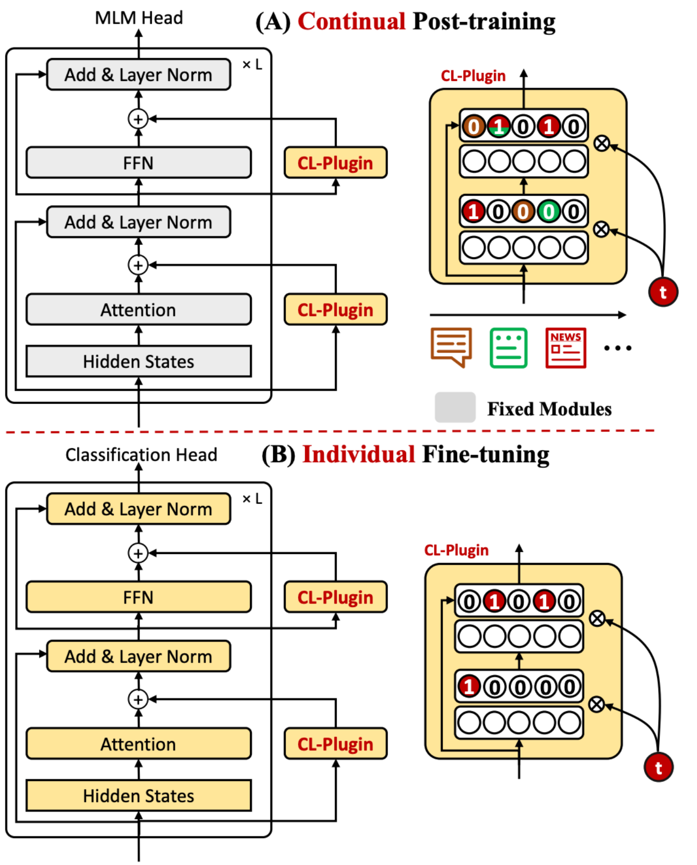
Overview

- From RLHF to Continual Learning
- A quick Introduction to Traditional Continual Learning
- Continual Learning of LMs
- Open Questions

What is special in CL for LM?

- LM is already pre-trained on some C0 (corpus 0)
  - C0 is usually unavailable – (CL challenge)
- Continual learning may be pre-training or fine-tuning
  - Pre-trained on C1 -> C2 -> C3 (domain adaptation, DIL w/o task-ID)
  - Fine-tuned on T1 -> T2 -> T3 (task adaptation, TIL or CIL)
- DIL
  - May happen in temporal dimension: evolution of language and knowledge
  - When it comes to open domain…
    - Alignment, model unlearning (learning to forget), model editing

Continual (DA-)pre-training



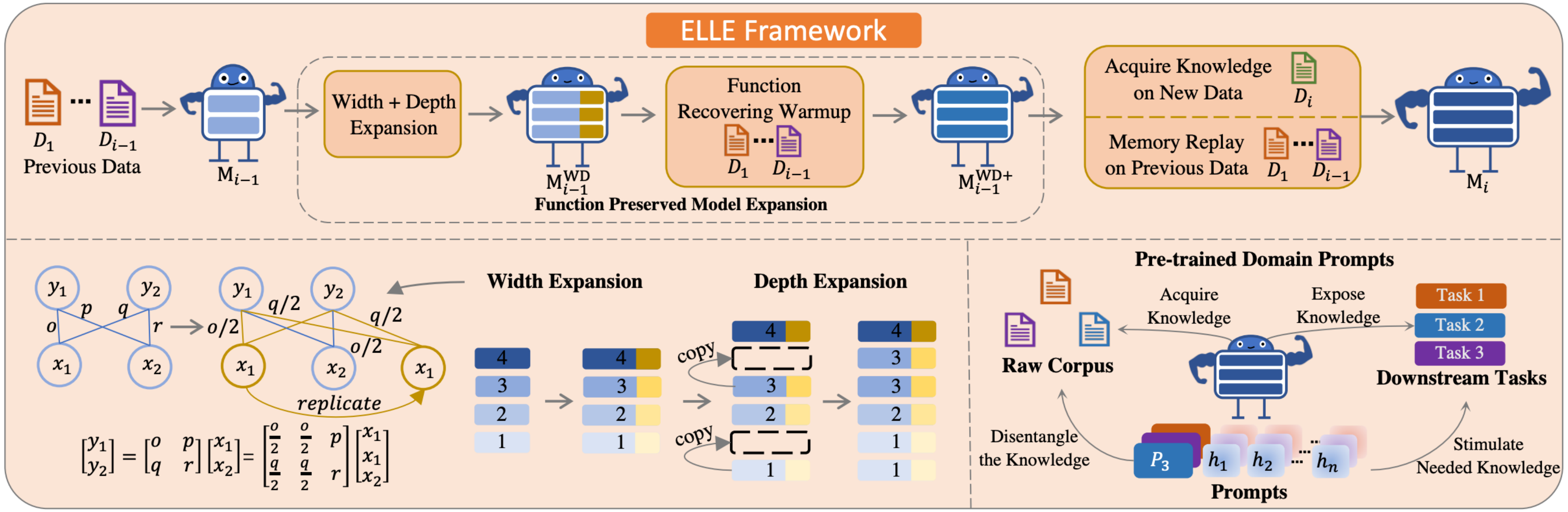
- train after pre-training: post-training / DAPT
  - Since DAPT helps downstream tasks, can we…
  - Language or knowledge may get outdated
  - generalist agent should be experts in multiple domains
- **Evaluation**: downstream fine-tuning performance
- **Baseline methods**
  - Naïve pre-training (+CF)
  - Parameter-isolation: Adapter, prompt, LoRA
  - Replay-based (memory should be large!)
  - CPT: hard-mask attention (+forward KT)

Category	Domain / Model	Restaurant MF1	Restaurant Acc	AI MF1	AI Acc	ACL MF1	ACL Acc	AGNews MF1	AGNews Acc	Average MF1	Average Acc	Forget R. MF1	Forget R. Acc
Non-CL	RoBERTa	50.61	74.77	27.88	28.44	32.19	34.59	64.19	65.95	43.72	50.94	—	
	Adapter	45.40	67.28	23.69	24.56	24.99	27.55	**64.53**	**66.50**	39.65	46.48	—	
	RoBERTa-ONE	53.63	**76.73**	29.86	30.11	33.05	35.72	62.57	65.13	44.78	51.92	—	
	Adapter-ONE	52.19	74.20	30.80	31.59	36.59	36.99	61.66	63.94	45.31	51.68	—	
	Prompt-ONE	28.93	59.79	21.06	22.10	28.02	29.22	60.70	62.58	34.68	43.42	—	
	DEMIX	53.14	75.28	27.68	27.29	37.63	38.57	63.18	65.13	45.41	51.57	—	
CL	RoBERTa-NCL	42.59	67.56	**31.57**	**31.62**	33.07	34.54	60.18	63.50	41.85	49.30	3.27	2.82
	Adapter-NCL	47.42	70.23	29.56	29.90	35.92	37.58	61.73	64.45	43.65	50.54	2.21	1.69
	HAT	50.45	71.78	28.33	29.41	34.93	37.15	62.97	65.05	44.17	50.85	2.43	2.04
	BCL	51.70	74.34	29.66	30.96	32.85	34.82	63.60	65.47	44.45	51.40	1.47	0.82
	KD	39.75	67.11	29.63	29.33	**38.30**	**42.09**	62.85	65.39	42.63	50.98	4.92	3.07
	EWC	48.32	71.59	30.96	31.01	35.96	38.05	62.29	64.95	44.38	51.40	1.40	0.80
	DER++	48.09	71.79	30.71	30.54	34.25	35.77	64.24	66.11	44.32	51.05	1.79	1.62
	CPT	**53.90**	75.13	30.42	30.89	37.56	38.53	63.77	65.79	**46.41**	**52.59**	**0.00**	**0.00**

Continual Training of Language Models for Few-Shot Learning, Zixuan Ke, Haowei Lin, Yijia Shao, et al. EMNLP (2022)

CPT literature (1)

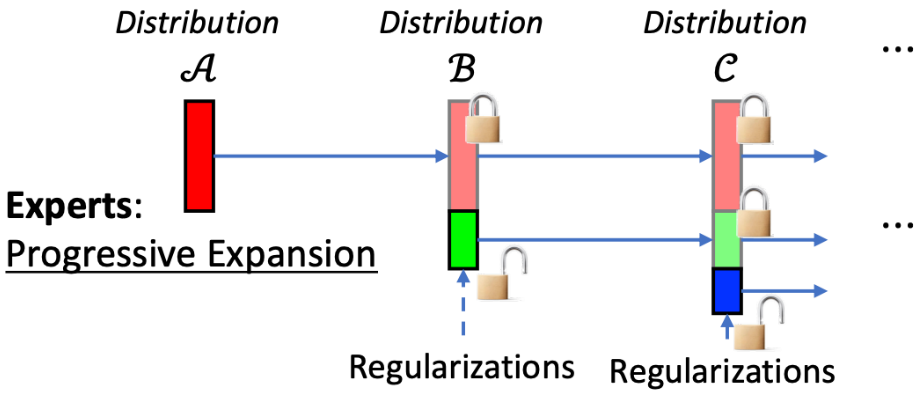
- **ELLE**
  - Network expansion + Replay + domain prompt (task-ID)

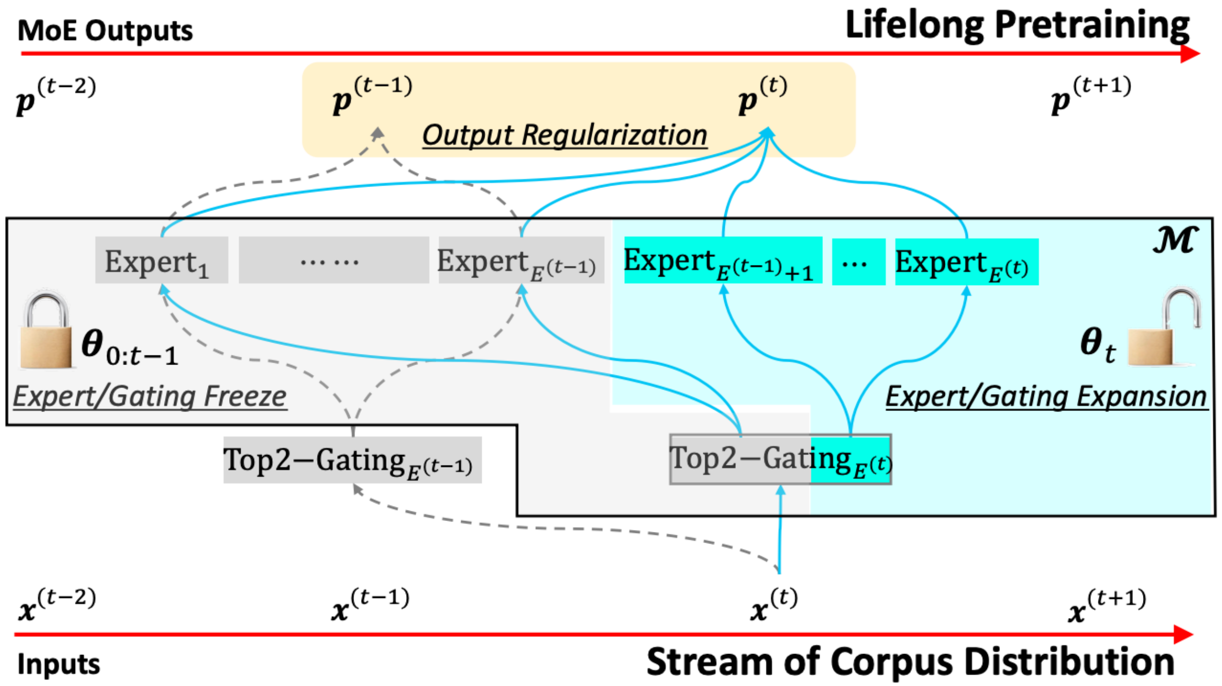


Qin, Yujia, et al. "ELLE: Efficient lifelong pre-training for emerging data." ACL 2022 findings

CPT literature (2)

- **Lifelong-MoE**
  - Regularization-based (distillation) + architecture-based

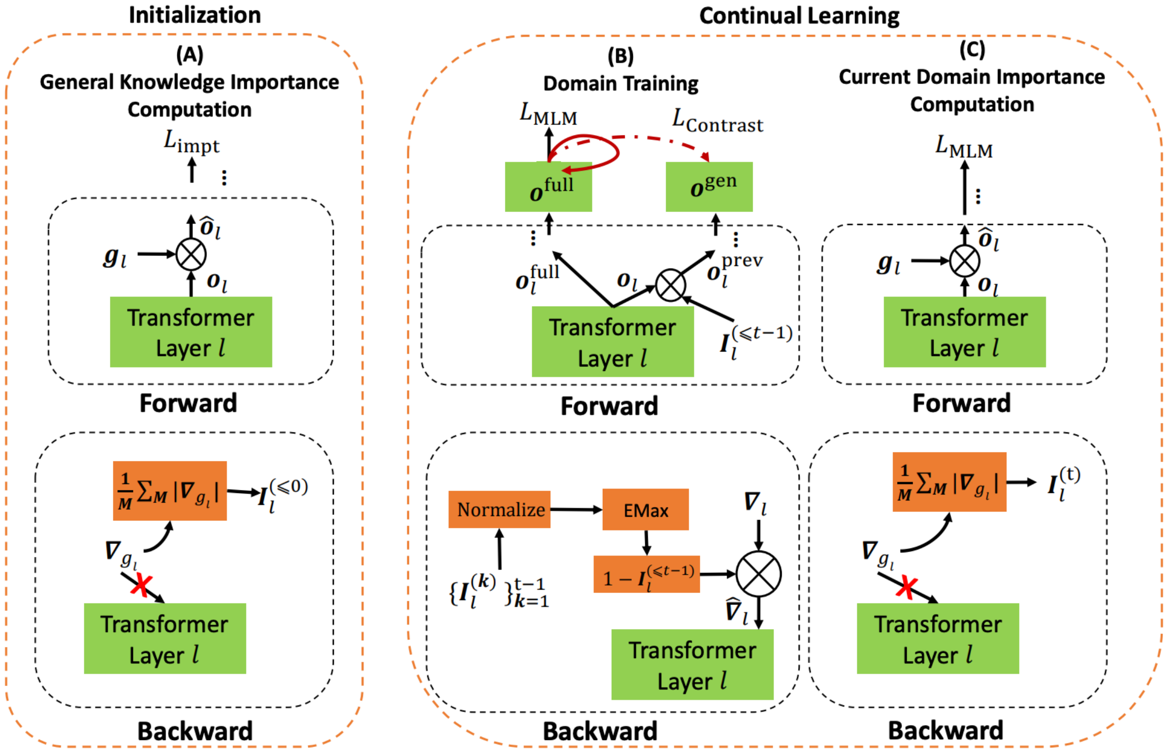




Chen, Wuyang, et al. "Lifelong Language Pretraining with Distribution-Specialized Experts." *International Conference on Machine Learning*. PMLR, 2023.

CPT literature (3)

- **DAS**: Continual DA-pre-training of LMs with Soft-masking)
  - Soft-masking (+forward & backward KT)



Zixuan Ke, Yijia Shao, Haowei Lin, et al. "Continual Pre-training of Language Models." *The Eleventh International Conference on Learning Representations*. 2022.

Findings

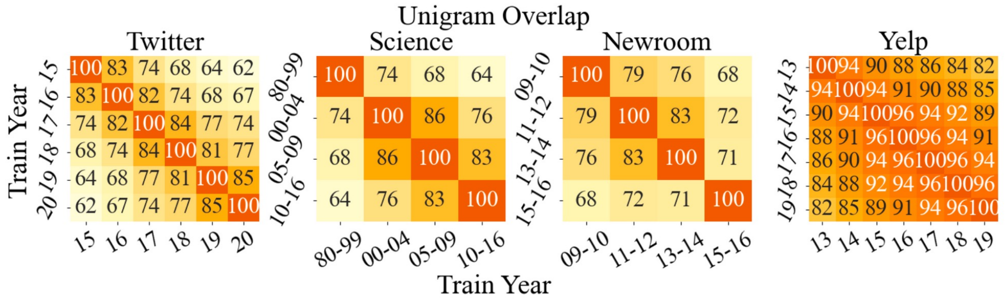
- Forgetting is minor
  - Cossu, Andrea, et al. "Continual pre-training mitigates forgetting in language and vision." arXiv preprint (2022).
  - An assumption: **generative loss is better than discriminative loss / small shift in both domain and task**
- CPT should be considered by GPT-5…? (when GPT-4 is outdated)
- Protection of general knowledge is crucial

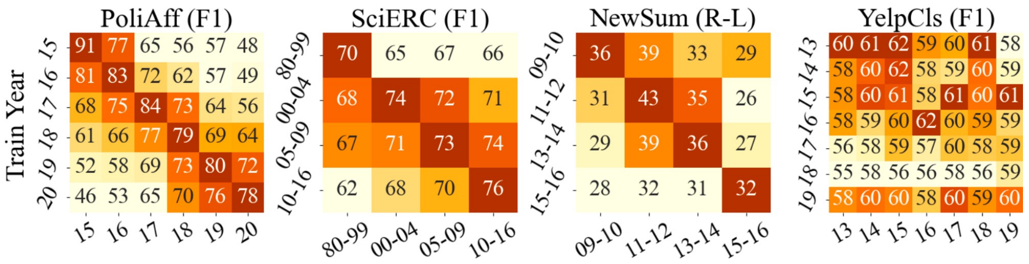
Domain	Camera		Phone		Resturant		AI		ACL		PubMed	Avg
Model	MF1	Acc.	MF1	Acc.	MF1	Acc.	MF1	Acc.	MF1	Acc.	Micro-F1	
RoBERTa	78.82	87.03	83.75	86.08	79.81	87.00	60.98	71.85	66.11	71.26	72.38	73.64
MLM	84.39	89.90	82.59	85.50	80.84	87.68	68.97	75.95	68.75	73.44	72.84	76.40
MLM (Adapter)	83.62	89.23	82.71	85.35	80.19	87.14	60.55	71.38	68.87	72.92	71.68	74.60
MLM (Prompt)	85.52	90.38	84.17	86.53	79.00	86.45	61.47	72.36	66.66	71.35	73.09	74.98
MLM+KD	82.79	89.30	80.08	83.33	80.40	87.25	67.76	75.46	68.19	72.73	72.35	75.26
MLM+AdaptedDeiT	86.86	91.37	83.08	85.64	79.70	86.84	69.72	76.83	69.11	73.35	72.69	76.86
MLM+SimCSE	84.91	90.35	83.46	86.08	80.88	87.59	69.10	76.25	69.89	74.30	72.77	76.84
MLM+TaCL	81.98	88.88	81.87	84.92	81.12	87.50	64.04	73.18	63.18	70.31	69.46	73.61
MLM+TaCO	84.50	90.22	82.63	85.32	79.27	86.68	59.73	71.22	63.66	70.36	72.38	73.69
MLM+InfoWord	87.95	91.92	84.58	86.84	81.24	87.82	68.29	75.92	68.58	73.68	73.21	77.31
DGA	**88.52**	**92.49**	**85.47**	**87.45**	**81.83**	**88.20**	**71.99**	**78.06**	**71.01**	**74.73**	**73.65**	**78.74**

Zixuan Ke, Yijia Shao, Haowei Lin, et al. “Adapting a Language Model While Preserving its General Knowledge.” EMNLP 2022.

Temporal Misalignment

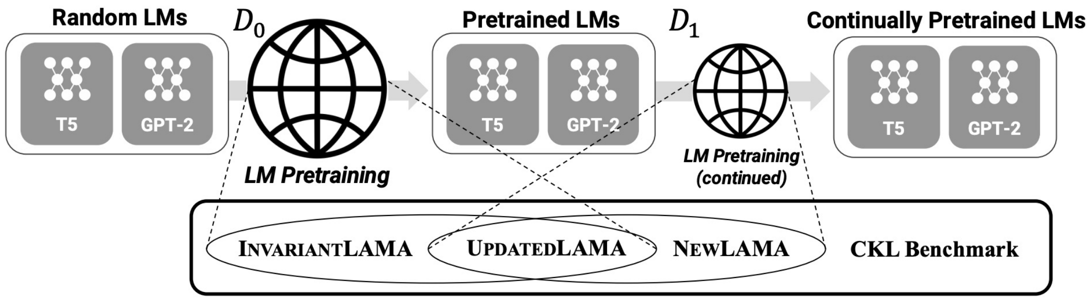
- **Temporal Misalignment** (TM)
  - training & evaluation datasets are from different periods of time
- RQs
  - How to assess TM?
  - How does TM affect downstream task performance?
  - The sensitivity to TM of different domains and tasks?
  - Can temporal adaptation (CPT) address TM?





Luu K, Khashabi D, Gururangan S, et al. Time waits for no one! analysis and challenges of temporal misalignment[J]. NAACL 2022.

Continual Knowledge Learning (CKL)



- Knowledge is dynamic
  - Retain time-invariant world knowledge
  - Update outdated knowledge
  - Acquire new knowledge
- Evaluation
  - LAMA
    - LAnguage Modeling Analysis
  - FUAR
    - [forgotten / (updated + acquired)]

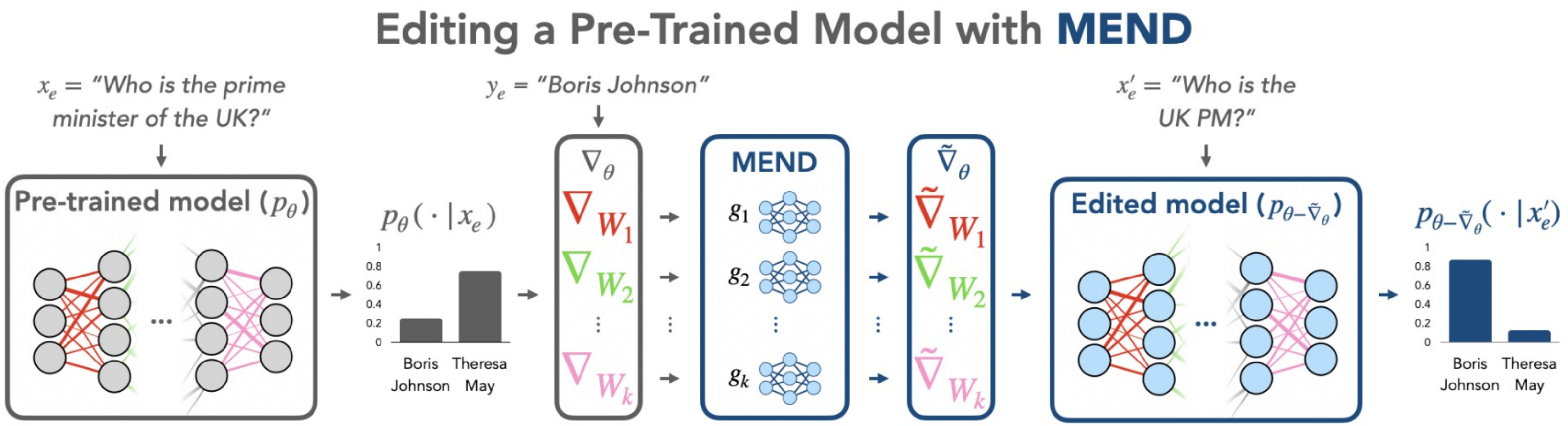
Method	# of Params (Trainable / Total)	IL EM	UL EM	NL EM	NLE EM	FUAR $((\mathbf{IL}), \mathbf{UL}, \mathbf{NL}) \downarrow$
T5-Initial	0M / 737M	**24.17**	1.62	1.88	10.32	-
T5-Vanilla	737M / 737M	12.89	10.17	3.77	17.75	1.08
T5-RecAdam	737M / 737M	13.20	12.55	4.02	17.85	0.84
T5-MixReview	737M / 737M	13.92	6.49	2.89	14.86	1.74
T5-LoRA	403M / 738M	16.58	**12.77**	4.52	**19.56**	0.55
T5-Kadapters (k=2)	427M / 762M	19.59	12.34	**5.03**	18.75	0.33
T5-Kadapters (k=3)	440M / 775M	19.76	12.66	4.02	19.00	0.33
T5-Modular	438M / 773M	20.29	12.66	4.65	19.24	**0.28**

Jang, Joel, et al. "Towards continual knowledge learning of language models." ICLR (2022).

Model Editing

Input	Pre-Edit Output	Edit Target	Post-Edit Output
1a: **Who is India's PM?**	Satya Pal Malik ✗	**Narendra Modi**	Narendra Modi ✓
1b: **Who is the prime minister of the UK?**	Theresa May ✗	**Boris Johnson**	Boris Johnson ✓
1c: Who is the prime minister of India?	Narendra Modi ✓	—	Narendra Modi ✓
1d: Who is the UK PM?	Theresa May ✗	—	Boris Johnson ✓
2a: **What is Messi's club team?**	Barcelona B ✗	**PSG**	PSG ✓
2b: **What basketball team does Lebron play on?**	Dallas Mavericks ✗	**the LA Lakers**	the LA Lakers ✓
2c: Where in the US is Raleigh?	a state in the South ✓	—	a state in the South ✓
3a: **Who is the president of Mexico?**	Enrique Pea Nieto ✗	**Andrés Manuel López Obrador**	Andrés Manuel López Obrador ✓
3b: Who is the vice president of Mexico?	Yadier Benjamin Ramos ✗	—	Andrés Manuel López Obrador ✗

- When LMs make errors / outdated…
  - A single problematic input vs. desired output is available
  - Fine-tuning tend to overfit
  - Tuning the whole model is computational infeasible or ineffective for LLMs
  - Similar to alignment (knowledge vs. safety)



Mitchell, Eric, Christopher D. Manning, et al. "Fast model editing at scale." NeurIPS (2022)

This is still mysterious!

Generative loss mitigates forgetting

- TIL and DIL will not be affected by CF much
  - But CIL still struggles with CF

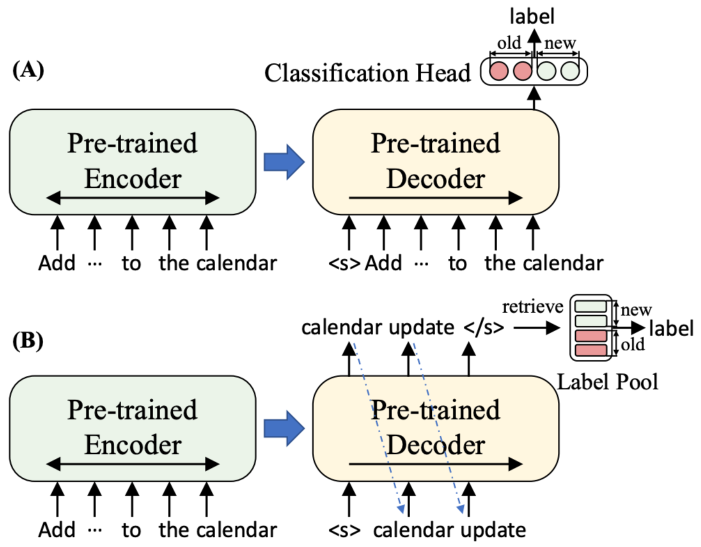


Figure 1: Comparison between classifier framework (A) and generation framework (B) of using a pre-trained encoder-decoder model for class-incremental learning.

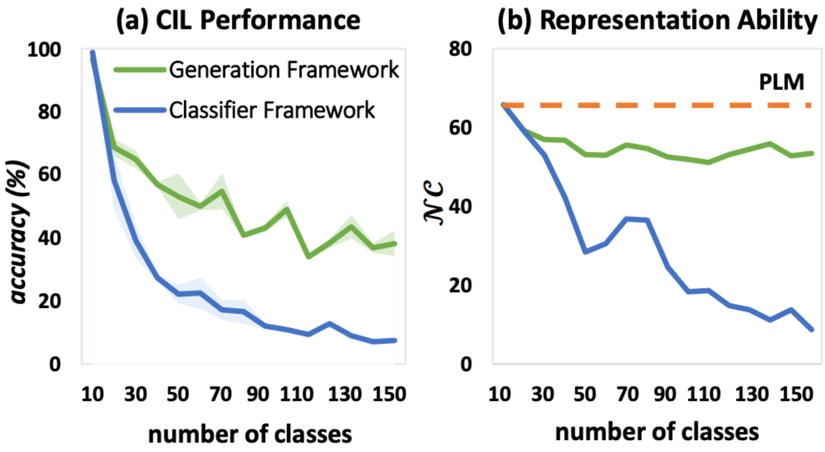


Figure 2: Accuracy (%) and $\mathcal{NC}$ (neural collapse) comparison of the classifier framework and generation framework for CIL on CLINC150 (15 tasks). For both *accuracy* and $\mathcal{NC}$, higher numbers are better.

Yijia Shao, et al. "Class-Incremental Learning based on Label Generation." *ACL* (2023).

Continual instruction tuning

Title	Answering simple science questions
Definition	In this subtask, you will answer a simple science question. Please indicate the correct answer. If you're not sure about the answer, choose the last option "I don't know".
Prompt	Please indicate the correct answer: A, B, C, D or E. If the question is not answerable or you're not sure about the answer, generate 'E' which implies "I don't know".
Positive example	**Input**: Question: When a guitar string is plucked, the sound is produced by (A) the size of the guitar. (B) the metal on the guitar. (C) the wood on the guitar. (D) the vibrations of the string. **Output**: D. **Explanation**: We know that the vibrations of the string produce sound in a guitar. So, the correct answer has to be "D".
Caution	The "A"-"D" responses correspond to the answer options mentioned in the input. There is a 5th option "E" which should be used for questions for which you're not sure about the answer (e.g., when the questions do not provide enough information to answer).
Things to avoid	Do not generate anything else apart from one of the following characters: 'A', 'B, 'C', 'D', 'E'.
Negative example	**Input**: A student found a rock while hiking in the mountains. By looking at the rock, she could tell the (A) exact weight of the rock. (B) length of time the rock had been on the hiking path. (C) color and shape of the rock. (D) exact length of the rock. **Output**: C i.e. color and shape of the rock. **Explanation**: "C" would have been a good answer. **Suggestions for fixing it**: You don't need to (and should not) explain the answer option.

Item	Explanation
Instruction-driven supervision	Each task is explained by an instruction and a couple of instances exemplifying it.
Fixed model capacity	The system's structure and parameter size are constant regardless of its learning status.
Knowledge maintenance	The system is not inclined to catastrophic forgetting.
Forward transfer	The system uses knowledge acquired from upstream tasks to help solve downstream tasks.
Backward transfer	The system uses knowledge acquired from downstream tasks to help solve upstream tasks.

Table 1: Desiderata of `ConTinTin`, inspired by (Biesialska et al., 2020).

Method		QG	AG	CF	IAG	MM	VF	mean
(Mishra et al., 2021)	paper report	52.xx	30.xx	50.xx	25.xx	47.xx	8.xx	35.33
	reimplement	53.55	17.45	63.79	11.06	82.86	7.40	39.35
Seq-finetune	forward	49.61	21.46	48.74	9.70	57.31	7.61	32.40
	backward	47.09	21.17	7.45	9.61	88.84	**14.98**	31.52
LAMOL	forward	52.23	20.45	67.74	8.81	82.29	8.83	40.05
	backward	52.14	22.76	7.98	8.33	88.45	9.91	31.59
`InstructionSpeak`	w/o CL	51.07	23.40	70.68	**11.43**	88.13	6.22	41.82
	forward	51.30	24.89	**70.96**	9.36	**90.41**	10.70	**42.93**
	backward	53.04	**24.93**	7.51	8.56	88.09	13.86	32.66

We don't see much forgetting on this generative task.
A good idea: from CIL to DIL (new formalization)

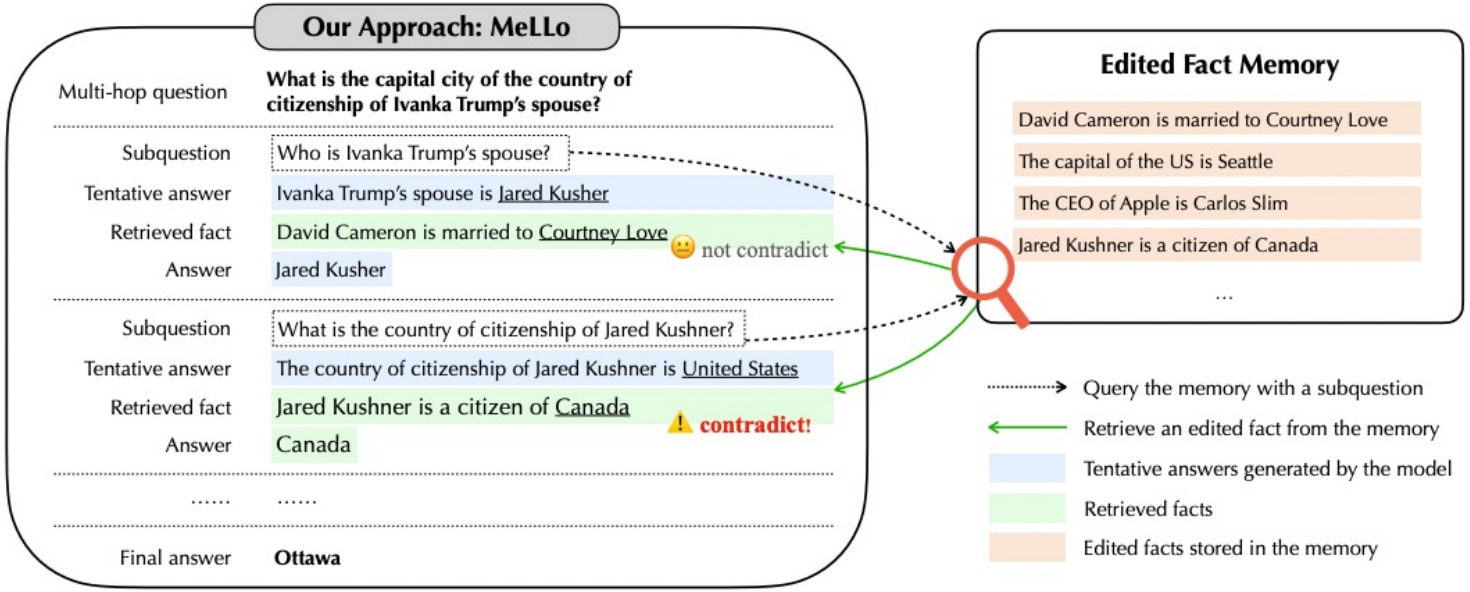
Yin, Wenpeng, Jia Li, and Caiming Xiong. "Contintin: Continual learning from task instructions." *ACL*(2022).

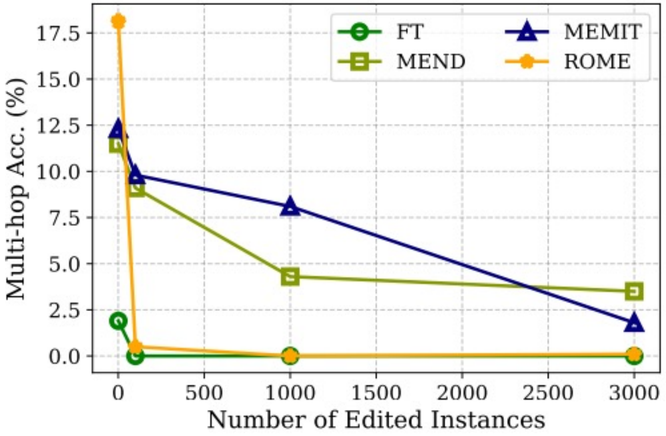
CL in the post-LLM Era

- **LLMs are infinity-task learners**
  - Traditional Classification-based TIL & CIL may be outdated (for building AGI)
  - Buzzy tasks, user-defined (creative) tasks, control (RL) tasks, multi-modality
- **Scaling, emergence, and reasoning**
  - Heated topics for LLMs, an it's still mysterious
  - they are missing in CL literature for many reasons
- **Memory-augmented LLMs**
  - A feasible choice for ML researchers to study continual learning
- **RLHF**
  - Continually learn from noisy human feedback

Memory-based model editing for LLMs

- **MeLLo** (Memory-based Editing for Large Language Models)
  - No training, scale to LLMs (w.r.t., MEND)
  - A new benchmark for multi-hop QA



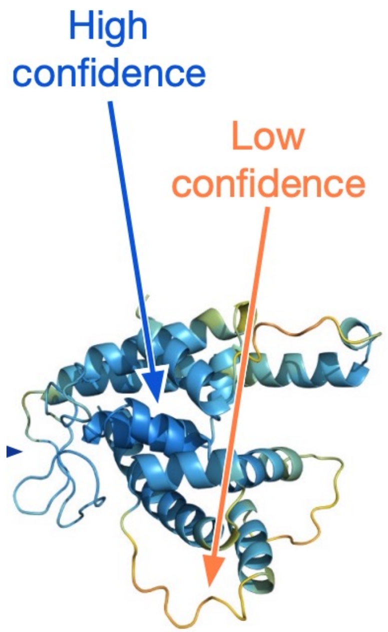


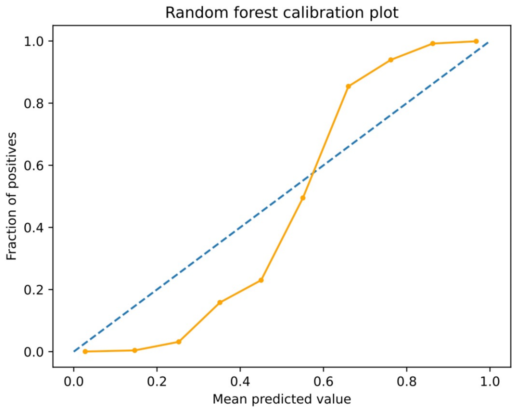
		# Edited instances			
Base Model	**Method**	1	100	1000	3000
GPT-J	MEMIT	12.3	9.8	8.1	1.8
GPT-J	MeLLo	20.3	12.5	10.4	9.8
Vicuna-7B	MeLLo	20.3	11.9	11.0	10.2
GPT-3	MeLLo	**68.7**	**50.5**	**43.6**	**41.2**

Zhong, Zexuan, et al. "MQuAKE: Assessing Knowledge Editing in Language Models via Multi-Hop Questions." arXiv preprint (2023).

LLMs have to know what they know

- **Out-of-distribution / anomaly / novelty detection**
  - Open-world learning (vs. close-world assumption)
  - Autonomy: Continually learn in an automatic way
  - Reject malicious noisy human feedback
  - Hallucination can be mitigated
  - Another very important topic related to CL
- Confidence learning
  - One of the most successful components in AlphaFold2
  - Difference in OOD detection: no ground truth
  - Another topic in ML community: model calibration





OOD detection establishes CIL SOTA

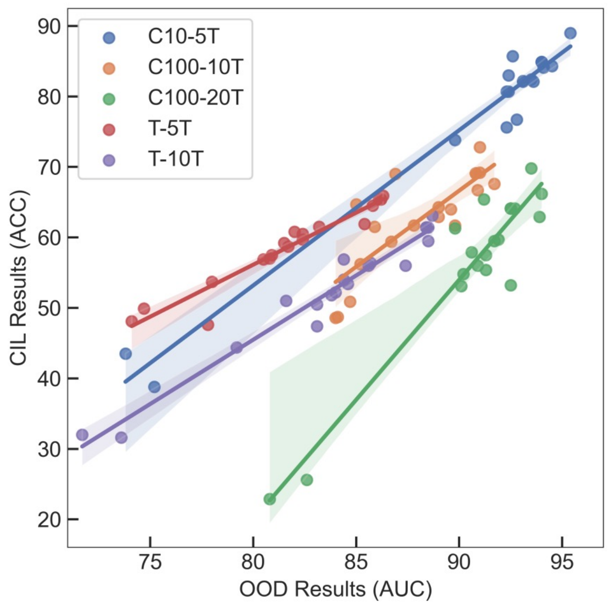


Figure 1: The correlation between OOD (AUC) and CIL (ACC) results. Each point denotes the AUC and ACC of one method in Tab. 1 on the same dataset.

	C10-5T	C100-10T	C100-20T	T-5T	T-10T	Average
Upper Bound	95.79$^{\pm0.15}$	82.76$^{\pm0.22}$	82.76$^{\pm0.22}$	72.52$^{\pm0.41}$	72.52$^{\pm0.41}$	83.70
OWM	41.69$^{\pm6.34}$	21.39$^{\pm3.18}$	16.98$^{\pm4.44}$	24.55$^{\pm2.48}$	17.52$^{\pm3.45}$	24.43
ADAM	83.92$^{\pm0.51}$	61.21$^{\pm0.36}$	58.99$^{\pm0.61}$	50.11$^{\pm0.46}$	49.68$^{\pm0.40}$	60.78
PASS	86.21$^{\pm1.10}$	68.90$^{\pm0.94}$	66.77$^{\pm1.18}$	61.03$^{\pm0.38}$	58.34$^{\pm0.42}$	68.25
HAT	82.40$^{\pm0.12}$	62.91$^{\pm0.24}$	59.54$^{\pm0.41}$	59.22$^{\pm0.10}$	54.03$^{\pm0.21}$	63.62
SLDA	88.64$^{\pm0.05}$	67.82$^{\pm0.05}$	67.80$^{\pm0.05}$	57.93$^{\pm0.05}$	57.93$^{\pm0.06}$	68.02
L2P	73.59$^{\pm4.15}$	61.72$^{\pm0.81}$	53.84$^{\pm1.59}$	59.12$^{\pm0.96}$	54.09$^{\pm1.14}$	60.47
iCaRL	87.55$^{\pm0.99}$	68.90$^{\pm0.47}$	69.15$^{\pm0.99}$	53.13$^{\pm1.04}$	51.88$^{\pm2.36}$	66.12
A-GEM	56.33$^{\pm7.77}$	25.21$^{\pm4.00}$	21.99$^{\pm4.01}$	30.53$^{\pm3.99}$	21.90$^{\pm5.52}$	31.19
EEIL	82.34$^{\pm3.13}$	68.08$^{\pm0.51}$	63.79$^{\pm0.66}$	53.34$^{\pm0.54}$	50.38$^{\pm0.97}$	63.59
GD	89.16$^{\pm0.53}$	64.36$^{\pm0.57}$	60.10$^{\pm0.74}$	53.01$^{\pm0.97}$	42.48$^{\pm2.53}$	61.82
DER++	84.63$^{\pm2.91}$	69.73$^{\pm0.99}$	70.03$^{\pm1.46}$	55.84$^{\pm2.21}$	54.20$^{\pm3.28}$	66.89
HAL	84.38$^{\pm2.70}$	67.17$^{\pm1.50}$	67.37$^{\pm1.45}$	52.80$^{\pm2.37}$	55.25$^{\pm3.60}$	65.39
MORE	89.16$^{\pm0.96}$	70.23$^{\pm2.27}$	70.53$^{\pm1.09}$	64.97$^{\pm1.28}$	63.06$^{\pm1.26}$	71.59
iFLP	**92.33**$^{\pm0.32}$	**76.53**$^{\pm0.27}$	**76.34**$^{\pm0.38}$	**68.64**$^{\pm0.44}$	**67.20**$^{\pm0.51}$	**76.21**

Haowei Lin, Yijia Shao, et al, “Class Incremental Learning by Exploiting OOD Data Distribution”, under review.

LMs know what they know

Language Models (Mostly) Know What They Know

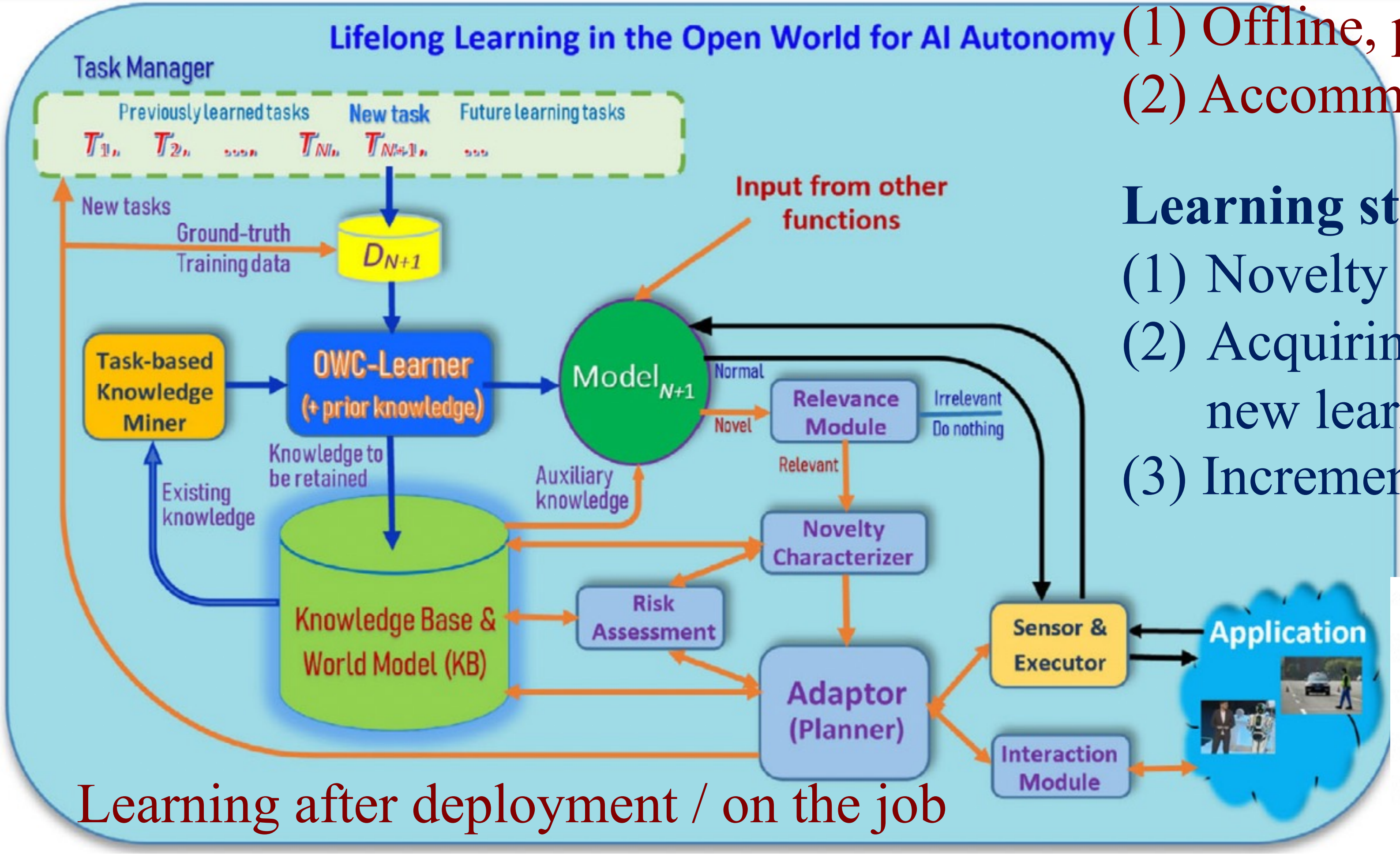
Saurav Kadavath,* Tom Conerly, Amanda Askell, Tom Henighan, Dawn Drain, Ethan Perez, Nicholas Schiefer, Zac Hatfield-Dodds, Nova DasSarma, Eli Tran-Johnson, Scott Johnston, Sheer El-Showk, Andy Jones, Nelson Elhage, Tristan Hume, Anna Chen, Yuntao Bai, Sam Bowman, Stanislav Fort, Deep Ganguli, Danny Hernandez, Josh Jacobson, Jackson Kernion, Shauna Kravec, Liane Lovitt, Kamal Ndousse, Catherine Olsson, Sam Ringer, Dario Amodei, Tom Brown, Jack Clark, Nicholas Joseph, Ben Mann, Sam McCandlish, Chris Olah, Jared Kaplan*

Anthropic

- Large models are well-calibrated on MC & QA
- RLHF policy miscalibration can be remediated
  - RL tends to collapse LM predictions towards behaviors with highest reward
  - Tuning with high temperature helps
- Self-evaluation
  - Similar to “Reflexion” (reflection)
- Limitations
  - Differentiate between “the truth” vs. “what human says”
  - Infinite recursion, generalization, etc.

- Kadavath et al. “Language Models (Mostly) Know What They Know” Arxiv 2022
- Shinn, Noah, Beck Labash, and Ashwin Gopinath. "Reflexion: an autonomous agent with dynamic memory and self-reflection." arXiv preprint arXiv:2303.11366 (2023).

*AI Autonomy: Self-initiated Open-world Continual Learning and Adaptation



Learning after deployment / on the job

(1) Offline, periodically -> self-initiated

(2) Accommodate to novel scenes

Learning steps

(1) Novelty detection

(2) Acquiring class labels and creating new learning tasks on the fly

(3) Incrementally learn the new task

User: Turn off the light in the kitchen
Bot: Sorry, I didn't get you. Do you mean to:
option-1. switch off the light in the kitchen,
option-2. switch on the light in the kitchen, or
option-3. change the color of the light?

Bing Liu, et al. AI Autonomy: Self-initiated Open-world Continual Learning and Adaptation. AI Magzine, 10 March 2023.

Recap

- Traditional CL
  - - CF, + KT, TIL, CIL, DIL, regularization, replay, architecture-based methods
- CL for LMs
  - DAPT & CPT, Temporal LM, Continual knowledge learning, model editing
  - Generative loss mitigates CF
    - (though CIL is unimportant, preserving general knowledge is still important)
- CL in the post-LLM era: from neural-based CL to system-based CL
  - Tasks become creative
  - Memory-based retrieval is promising
  - OOD detection, confidence / calibration

Q&A?

Happy for further discussion:

linhaowei@pku.edu.cn

linhaowei1.github.io