



大连理工大学

信息检索研究室

Information Retrieval Laboratory of DUT

大模型与知识蒸馏技术

Knowledge Distillation (KD)

调研负责：王宏博

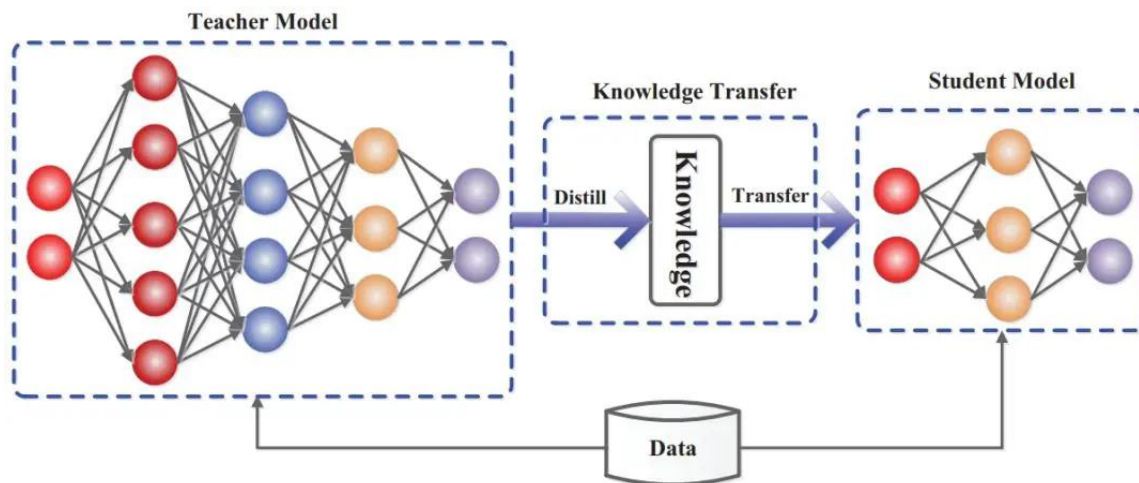
- 一、引言
- 二、传统时代的KD
- 二、大模型时代的KD
- 三、如何在研究中应用



引言

● 知识蒸馏

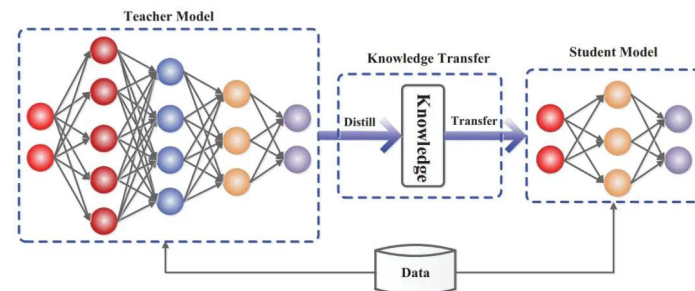
- 标准定义：“知识蒸馏技术”是使用一个**大型**的“教师模型 (Teacher Model)”来指导一个**小型**的“学生模型 (Student Model)”的训练
- 学生模型不再从头用巨量数据预训练，而直接学习教师模型的“既定经验”，从而快速掌握教师的理解能力



● 为什么使用知识蒸馏？

- 节约成本：麻雀虽小，五脏俱全
- 更强的泛化能力
- 闭源-开源优势（GPT4- > Orca）

...

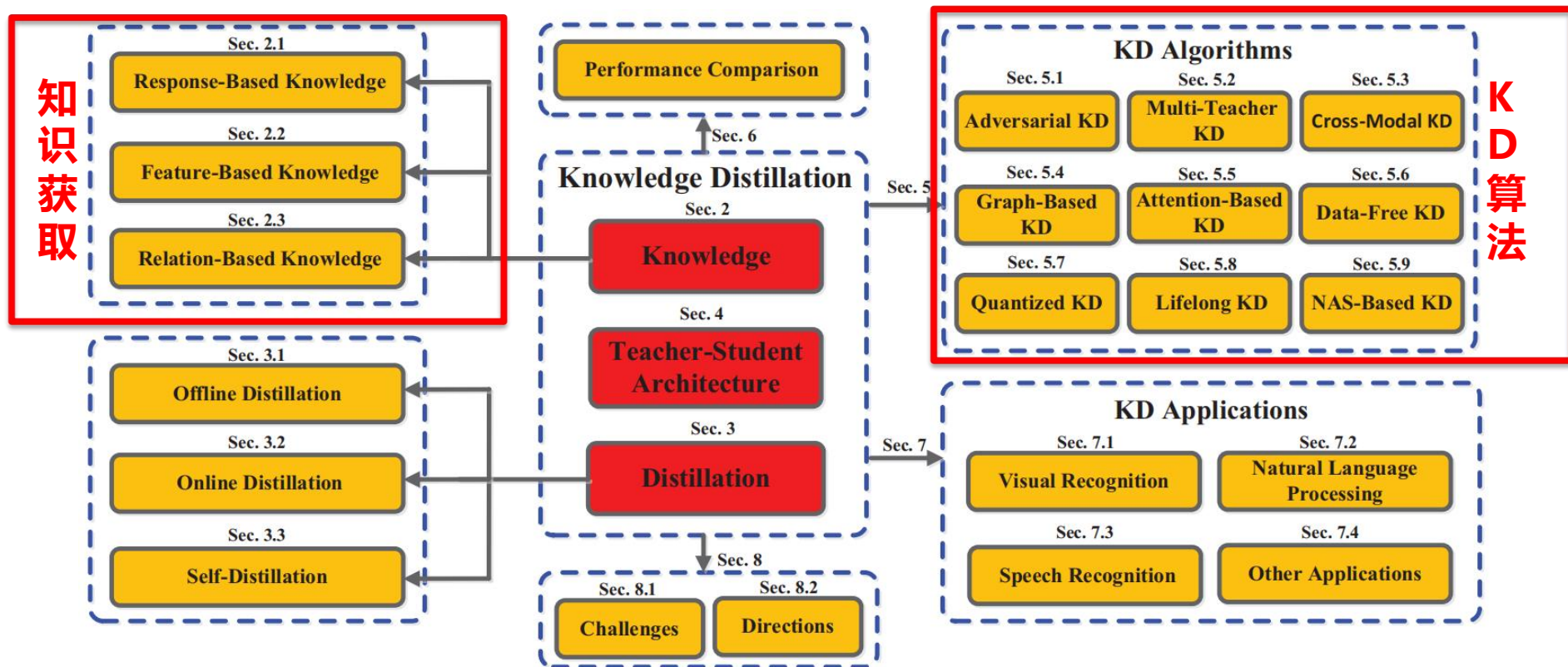




传统时代的KD

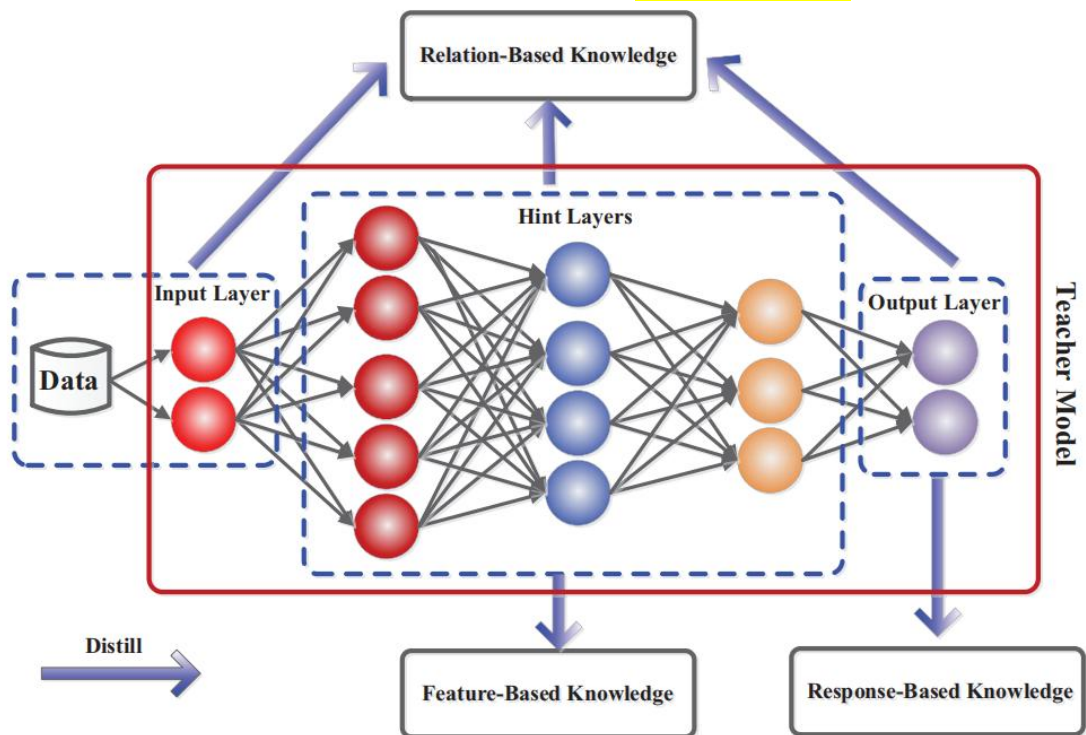
● 知识蒸馏流程(传统)

➤ 知识获取 -> KD算法知识蒸馏 -> 蒸馏应用



● 知识获取(传统)

- 为了学生模型的学习，我们需要获取教师模型的“既定经验”
- 传统的知识蒸馏侧重于“架构”的模仿



特征式知识

Feature-Based

响应式知识

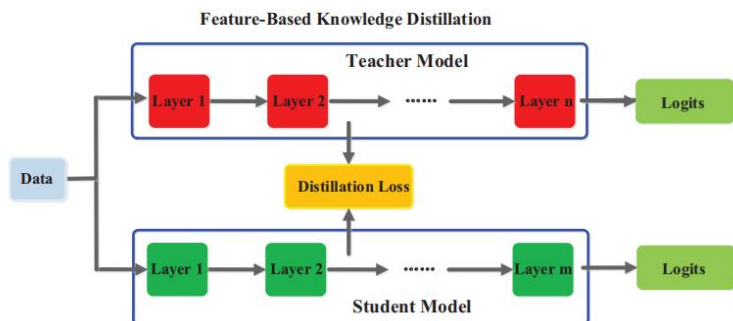
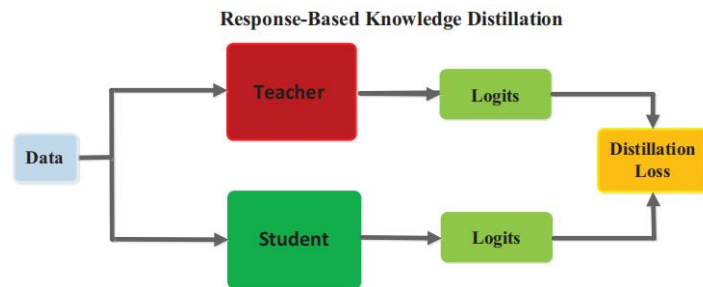
Response-Based

关联式知识

Relation-Based

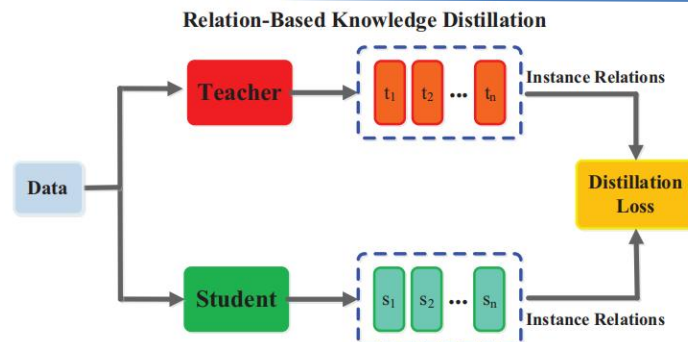
知识获取(传统)

$$L_{ResD}(p(z_t, T), p(z_s, T)) = \mathcal{L}_R(p(z_t, T), p(z_s, T))$$



$$L_{FeaD}(f_t(x), f_s(x)) = \mathcal{L}_F(\Phi_t(f_t(x)), \Phi_s(f_s(x)))$$

$$L_{RelD}(F_t, F_s) = \mathcal{L}_{R^2}(\psi_t(t_i, t_j), \psi_s(s_i, s_j))$$

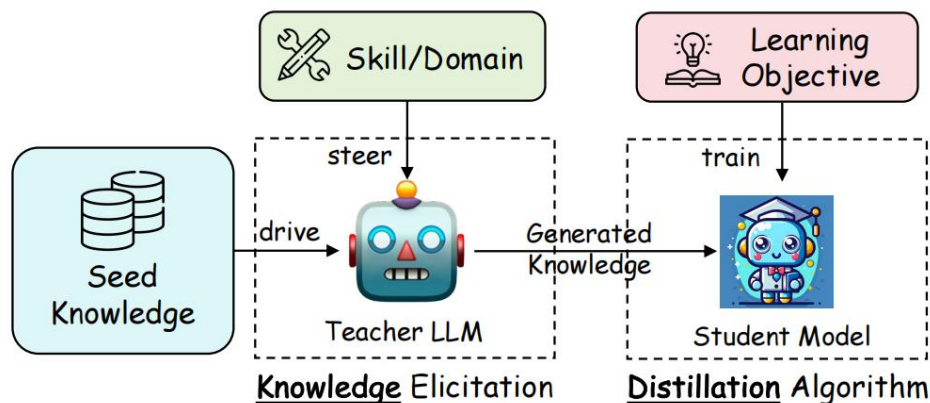




大模型时代的KD

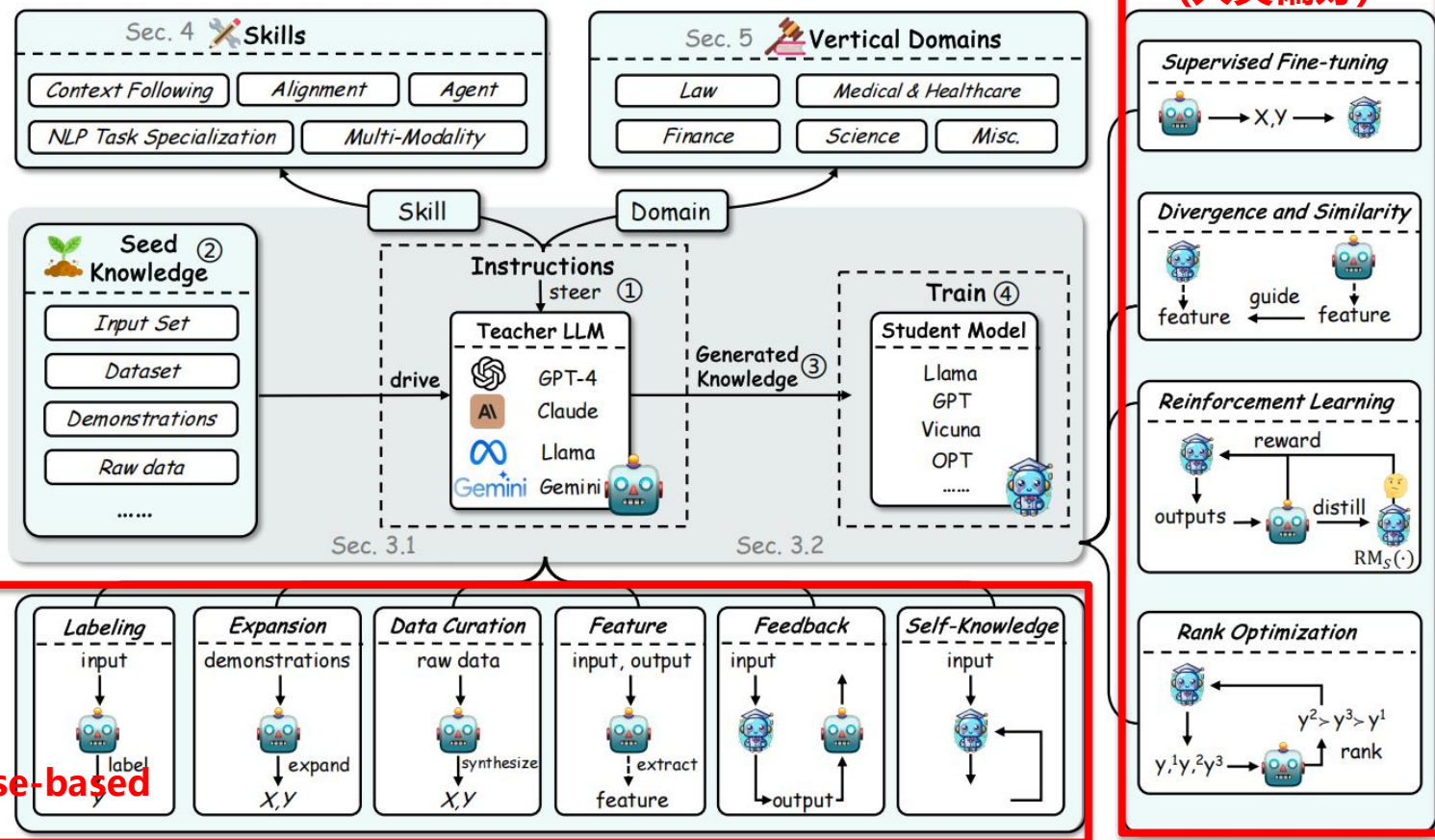
● 知识蒸馏流程(LLM)

- 目的一致，仍是为了学生模型 (Open-Source LLMs) 获取教师模型(Closed/Open-Source LLMs)的“既定经验”，并加以学习
- **重要区别：** 1. 知识获取侧重于Response-based Knowledge，也就是“数据知识”的模仿 --- **成本？ 闭源？ 效果？**
2. KD算法侧重于SFT+RL偏好，以适应生成式模型



● 知识蒸馏流程(LLM)

➤ 知识获取 -> KD算法知识蒸馏 -> 蒸馏应用



知识
获取

Response-based

● 知识蒸馏流程(LLM)

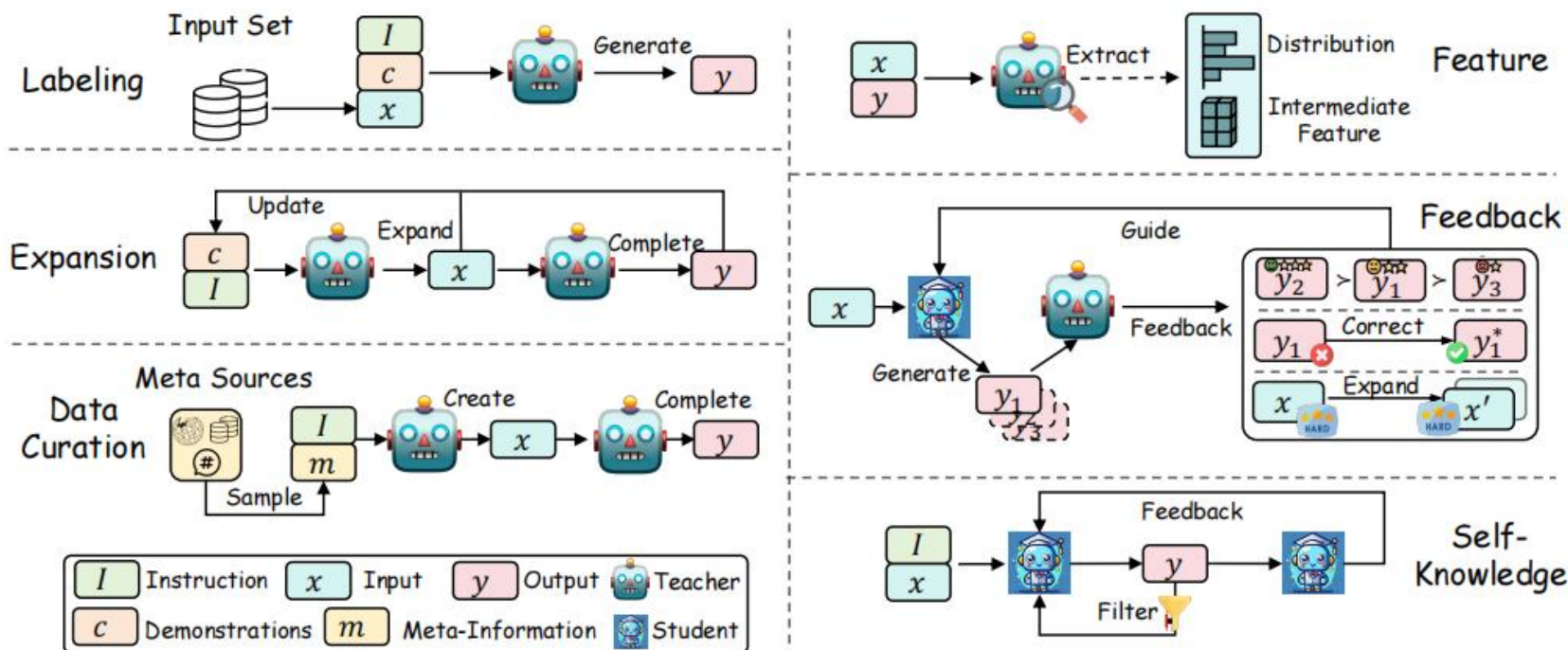
➤ 知识获取: I 指令 S 种子 $\mathcal{D}_I^{(kd)}$ 蒸馏知识

$$\mathcal{D}_I^{(kd)} = \{\text{Parse}(o, s) | o \sim p_T(o | I \oplus s), \forall s \sim \mathcal{S}\} \quad (1)$$

➤ KD算法: 多任务 \sum_I 学生模型参数 θ_S

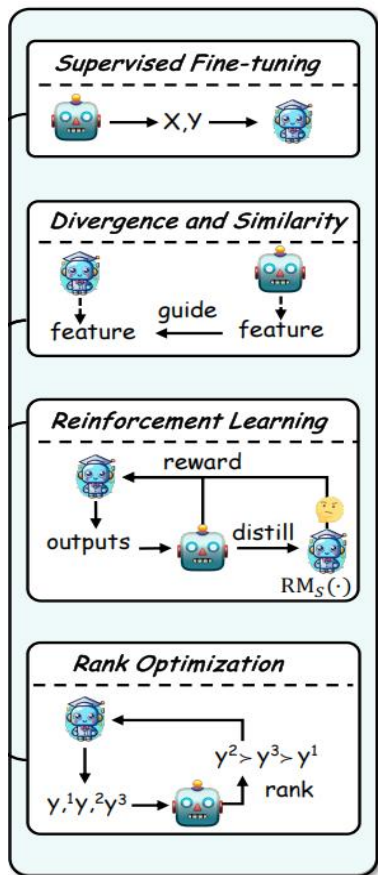
$$\mathcal{L} = \sum_I \mathcal{L}_I(\mathcal{D}_I^{(kd)}; \theta_S) \quad (2)$$

● 知识获取(LLM)



- LLM的知识获取:
标注(Labeling)、扩展(Expansion)、数据整理(Data Curation)、
特征提取 (Feature)、反馈(Feedback)、自我知识(Self-Knowledge)

● KD算法(LLM) -- 类似于LLM的相关算法，来源不同



➤ Supervised Fine-Tuning (Instruction Distillation)

也称为指令蒸馏算法，是目前最通用的黑盒LLM KD

➤ Divergence and Similarity

用于经由隐藏层生成的Feature的模型指导，适用于白盒LLM

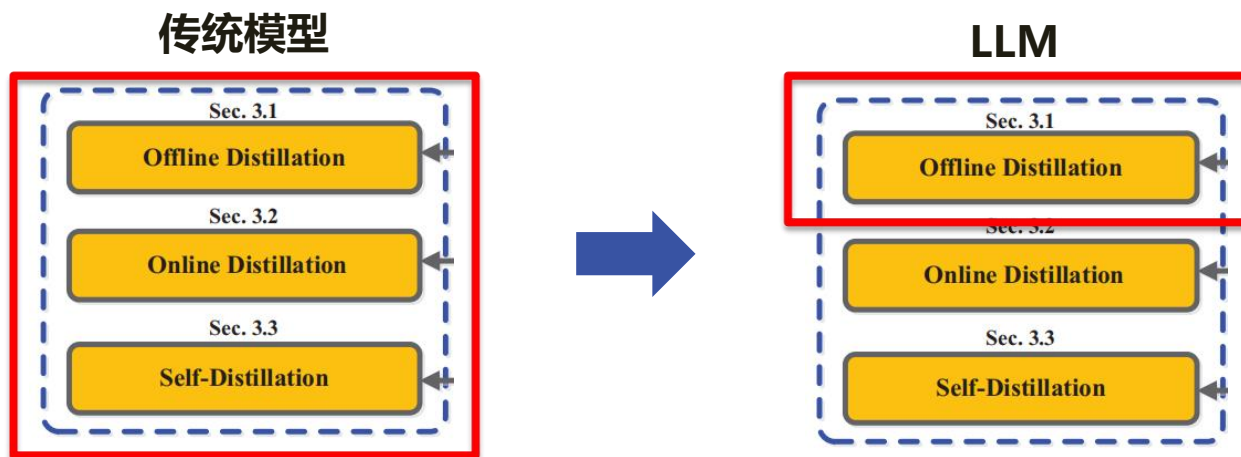
➤ Reinforcement Learning

教师评价学生的输出，奖励符合教师标准的输出

➤ Rank Optimization

优化后的RL算法，更加高效（如DPO）

● 知识蒸馏流程(LLM)



由于教师模型参数层很少参与蒸馏过程，知识先被提取（数据集）再进行KD，因此LLM的KD更多采用**离线蒸馏-Offline Distillation**



如何在研究中的应用

- 我们的目的是高效的将大模型的知识**集成**给小模型，节约成本

➤ 我们需要：

1. 获得这些知识（自建数据集/开源的类官方数据）
2. 将知识进一步集成于我们现有的大模型（如太学，太令等...），以及现有研究，将它们作为学生模型
3. 集成方向：
 - > **Context Following** 上下文跟随
 - > **Alignment** 思维范式，偏好对齐
 - > **Agent** 擅长工具的代理
 - > **NLP Task Specialization** NLU/NLG/IR

● 上下文跟随能力 - Context Following

- 该过程增强学生LLMs处理各种复杂上下文的能力，例如少样本示例、复杂指令、多轮对话以及检索增强信息

Methods	Skill	Seed Knowledge	Teacher LLM	Student Model	Knowledge Elicitation	Objective
Context Following						
Self-Instruct (Wang et al., 2022a)	IF	175 human-curated tasks	GPT3	GPT3	Expansion + Self-Knowledge	SFT
Alpaca (Taori et al., 2023)	IF	175 human-curated tasks	GPT3	LLaMA	Expansion + Self-Knowledge	SFT
LaMini-LM (Wu et al., 2023c)	IF	3.5K Wikipedia Categories + Mixed Dataset	ChatGPT	Various Models	Expansion	SFT
WizardLM (Xu et al., 2023a)	IF	Alpaca Data	ChatGPT	LLaMA	Expansion	SFT
Lion (Jiang et al., 2023b)	IF	Alpaca Cata	ChatGPT	LLaMA	Labeling + Expansion + Feedback	-
BabyLlama (Timiryasov and Tastet, 2023)	IF	10M-word BabyLM dataset	GPT-2 + small LLaMA	58M-parameter LLaMA	Feature	D&S
MiniLLM (Gu et al., 2024)	IF	Dolly Dataset	GPT2 + OPT + LLaMA	GPT2 + OPT + LLaMA	Feature	D&S
Self-Align (Sun et al., 2024b)	IF	Human-written Principles	LLaMA	LLaMA	Expansion + Self-Knowledge	SFT
Self-Rewarding (Yuan et al., 2024a)	IF	Human-written Samples	LLaMA	LLaMA	Self-Knowledge	SFT + RL
STaR (Zelikman et al., 2022)	IF	Arithmetic + CommonsenseQA + GSM8K	GPT-J	GPT-J	Self-Knowledge	SFT
Llama-GPT4 (Peng et al., 2023a)	IF	Alpaca Dataset	GPT4	LLaMA	Labeling	SFT
Reflection-Tuning (Li et al., 2023e)	IF	Alpaca/WizardLM Dataset	ChatGPT	LLaMA	Labeling	SFT
Selective Reflection-Tuning (Li et al., 2024d)	IF	Alpaca/WizardLM Dataset	ChatGPT	LLaMA	Labeling	SFT
Vicuna (Chiang et al., 2023)	IF/MD	Human Conversation	ChatGPT + GPT4	LLaMA	Labeling	SFT
Koala (Geng et al., 2023)	IF/MD	Human Conversation	ChatGPT	LLaMA	Labeling	SFT
Baize (Xu et al., 2023b)	IF/MD	Quora + Stack Overflow	ChatGPT	LLaMA	Expansion + Self-Knowledge	SFT
UltraChat (Ding et al., 2023b)	IF/MD	Wikidata + Text Material + C4	ChatGPT	LLaMA	Curation	SFT
Orca (Mukherjee et al., 2023)	IF/TP	FLAN-v2	ChatGPT + GPT4	LLaMA	Labeling	SFT
Orca2 (Mittra et al., 2023)	IF/TP	FLAN-v2 + Few-Shot/Math/Synthetic	GPT4	LLaMA	Labeling	SFT
SelfFree (Ye et al., 2023)	IF/TP	Human Conv, Flan/Code/Math Collection	ChatGPT	LLaMA	Labeling	SFT
CoT-Distill (Hsieh et al., 2023)	IF/TP	e-SNLI + ANLI + CQA + SVAMP	PaLM	T5	Labeling	SFT
KnowPAT (Zhang et al., 2023a)	IF/TP	CPKG + QA Data	ChatGPT + ChatGLM + Vicuna-7B	LLaMA	Labeling	SFT
DEBATunE (Li et al., 2024e)	IF/TP	Controversial Topics	ChatGPT	LLaMA	Labeling	SFT
Phi-1 (Gunasekar et al., 2023)	IF/Code	-	GPT3.5	phi-1	Curation	SFT
Phi-1.5 (Li et al., 2023a)	IF/Code	20k Topics from Web	GPT3.5	phi-1	Curation + Labeling	SFT
SAIL (Luo et al., 2023c)	IF/RAG	Alpaca Data + Web Content	GPT4	LLaMA	Label	SFT
KARD (Kang et al., 2023b)	IF/RAG	MedQAUSMLE	ChatGPT	T5 + OPT	Label	SFT + D&S
Self-RAG (Asai et al., 2023)	IF/RAG	Open-Instruct	GPT4	LLaMA	Labeling	SFT

● 对齐能力 - Alignment

- 该过程使学生的输出与教师在语义、风格、偏好和价值观念等方面保持一致，确保其生成的内容符合人类的期望和伦理标准

Methods	Skill	Seed Knowledge	Teacher LLM	Student Model	Knowledge Elicitation	Objective
<i>Alignment</i>						
OpenChat (Wang et al., 2023c)	IF/Preference	Human Conversation	ChatGPT + GPT4	LLaMA	Labeling	SFT + RL
Zephyr (Tunstall et al., 2023)	IF/Preference	Mixed Datasets	GPT4	Mistral	Labeling + Feedback	SFT + RO
ALMoST (Kim et al., 2023a)	IF/Preference	Human-written Prompts	LLaMA	LLaMA	Expansion + Labeling	SFT + RL
RLCD (Yang et al., 2024)	IF/Preference	Human-written Prompts	LLaMA	LLaMA	Labeling	SFT + RL
RLAIF (Lee et al., 2023a)	IF/Preference	Human-written Prompts	PaLM 2	PaLM 2	Labeling + Feedback	RL
GPT3 Reward (Kwon et al., 2023)	Preference	Human-written Prompts	GPT3	GPT3	Labeling	RL
ILF (Scheurer et al., 2023)	Preference	Task-specific Datasets	GPT3 + FeedME	GPT3	Labeling	RL
ULTRAfeedback (Cui et al., 2023a)	Preference	Mixed Datasets	GPT4	LLaMA	Labeling	RL
Constitutional AI (Bai et al., 2022a)	Preference/Value	Human-written Prompts	Self-defined Student Model	Self-defined Model	Labeling + Expansion + Feedback	SFT + RL
SANDBOX (Liu et al., 2023b)	Value	Simulation	text-davinci-002/-003 + GPT4 + ChatGPT	LLaMA	Data Curation	SFT + RL

● 自主能力 - Agent

- 该过程增强学生LLMs的工具使用和规划能力，使其能够像人类智能体一样高效地执行复杂任务

Methods	Skill	Seed Knowledge	Teacher LLM	Student Model	Knowledge Elicitation	Objective
<i>Agent</i>						
Toolformer (Schick et al., 2023)	Tool	CCNet	GPT-J	GPT-J	Labeling	SFT
Graph-ToolFormer (Zhang, 2023)	Tool	Mixed Graph Dataset	ChatGPT	GPT-J + LLaMA	Labeling	SFT
Gorilla (Patil et al., 2023)	Tool	Online API Documentation	GPT4	LLaMA	Expansion	SFT
GPT4Tools (Yang et al., 2023b)	Tool	Image Content	ChatGPT	LLaMA	Curation + Expansion	SFT
ToolAlpaca (Tang et al., 2023a)	Tool	Public-apis Repository	ChatGPT	LLaMA	Curation	SFT
ToolLLM (Qin et al., 2023a)	Tool	Real-world APIs	ChatGPT	LLaMA	Curation	SFT
MLLM-Tool (Wang et al., 2024)	Tool	HuggingFace Model Cards	GPT4	LLaMA	Curation	SFT
FireAct (Chen et al., 2023b)	Planning	Mixed QA Dataset	GPT4	LLaMA	Labeling	SFT
AgentTuning (Zeng et al., 2023a)	Planning	6 Agent Tasks	GPT4 + ChatGPT	LLaMA	Labeling + Expansion	SFT
Lumos (Yin et al., 2023a)	Planning	Mixed Interactive Tasks	GPT4	LLaMA	Labeling	SFT
AUTOACT (Qiao et al., 2024)	Planning	Mixed QA Tasks	LLaMA	LLaMA	Labeling	SFT

● 领域泛化能力 - NLP Task Specialization

➤ 使学生模型在特定自然语言处理任务上表现更佳，提升其在不同NLP领域的适应性和性能（尤其是MLLM）

Methods	Skill	Seed Knowledge	Teacher LLM	Student Model	Knowledge Elicitation	Objective
<i>NLP Task Specialization</i>						
AugGPT (Dai et al., 2023a)	NLU	Amazon/Symptoms/PubMed20k Dataset	ChatGPT	BERT	Label	SFT
TDG (He et al., 2023b)	NLU	SST + QQP + MNLI	GPT3	BERT	Expansion	SFT
SunGen (Gao et al., 2023a)	NLU	Text Classification Tasks	GPT2	DistilBERT	Curation	SFT
UDG (Wang et al., 2021a)	NLU	NLU Tasks	GPT3	BERT	Expansion	SFT
InheritSumm (Xu et al., 2023c)	NLG	Pile + ArXiv + CNN/DM + WikiHow	GPT3.5	ZCode++	Label	SFT
DIMSUM+ (Jung et al., 2023)	NLG	None	GPT2 + CTRL + BioGPT	T5	Curation + Self-Knowledge	SFT
Genie (Yehudai et al., 2024)	NLG	ELI5 + ASQA + NQ + CNN/DM	Falcon + LLaMA	FLAN + LLaMA	Label	SFT
GKD (Agarwal et al., 2024)	NLG/NLU/IF	XSum+WMT14 en-de+GSM8K+FLAN2021	T5-XL	T5	Feature + Feedback	D&S + RL
QUILL (Srinivasan et al., 2022)	IR	IR Datasets	T5	4-layer Transformer	Internal Knowledge	D&S
RankVicuna (Pradeep et al., 2023a)	IR	IR Datasets	ChatGPT	LLaMA	Labeling	SFT
RankZephyr (Pradeep et al., 2023b)	IR	IR Datasets	ChatGPT + GPT4	Mistral	Labeling	SFT
NDR (Mysore et al., 2023)	Recommendation	Recommendation Datasets	GPT3	MPnet-110M	Labeling	SFT
InstructRec (Zhang et al., 2023b)	Recommendation	39 instruction templates	ChatGPT	Flan-T5	Expansion + Self-Knowledge	SFT
ONCE (Liu et al., 2023c)	Recommendation	Recommendation Dataset	ChatGPT	LLaMA	Labeling	SFT
PandaLM (Wang et al., 2023b)	Evaluation	Alpaca Data	ChatGPT	LLaMA	Labeling	SFT
Prometheus (Kim et al., 2024)	Evaluation	50 Seed Rubrics	GPT4	LLaMA	Labeling	SFT
InstructScore (Xu et al., 2023d)	Evaluation	Mixed Dataset	GPT4	LLaMA	Labeling	SFT
WizardMath (Luo et al., 2023b)	Math	GSM8k + MATH	ChatGPT	LLaMA	Expansion + Feedback	SFT + RL
Mammoth (Yue et al., 2023a)	Math/TP	Mixed Math Dataset	GPT4	LLaMA	Labeling	SFT
Mixed Distill (Chenglin et al., 2023)	Math/TP	SVAMP + GSM8K + ASDIV + StrategyQA	ChatGPT	LLaMa	Labeling	SFT
WizardCoder (Luo et al., 2023a)	Code	Code Alpaca Data	ChatGPT	StarCoder	Expansion	SFT
MagiCoder (Wei et al., 2023)	Code	Existing Source Codes	ChatGPT	LLaMa	Curation	SFT
WaveCoder (Yu et al., 2024)	Code	Existing Source Codes	GPT4	LLaMa	Curation	SFT
Code Alpaca (Chaudhary, 2023)	Code	Code Instructions	ChatGPT	LLaMA	Expansion + Self-Knowledge	SFT
Code Llama (Rozière et al., 2023)	Code	Human-written Instructions	LLaMA	LLaMA	Expansion + Self-Knowledge	SFT
Code Clean (Jain et al., 2023)	Code	Code Datasets	ChatGPT	LLaMA	Labeling	SFT
<i>Multi-Modality</i>						
LLaVA (Liu et al., 2023e)	Vision-Language	COCO	GPT4	LLaMA	Labeling	SFT
SVIT (Zhao et al., 2023b)	Vision-Language	Visual Genome + COCO	GPT4	LLaMA	Labeling	SFT
LVIS-Instruct4V (Wang et al., 2023e)	Vision-Language	LVIS	GPT4V	LLaMA	Labeling	SFT
LLaVAR (Zhang et al., 2023d)	Vision-Language	LAION	GPT4	LLaMA	Labeling	SFT
Macaw-LLM (Lyu et al., 2023)	Multiple Modalities	Image/Video with Caption	ChatGPT	LLaMA	Labeling	SFT
MIMIC-IT (Li et al., 2023f)	Multiple Modalities	Image/Video Dataset	ChatGPT	LLaMA	Labeling	SFT
ChatBridge (Zhao et al., 2023d)	Multiple Modalities	Task-Specific/Multimodal-Chat Data	GPT4 + ChatGPT	LLaMA	Labeling	SFT

● 应用实例- Deepseek的蒸馏与复现

- Deepseek已经证明，通过知识蒸馏，将高级推理能力迁移到较小的模型中是可行且实用的
- DeepSeek-R1-Distill是使用DeepSeek-R1 蒸馏的样本对开源模型如Qwen/llama微调得到的，**规模小，成本低，效果好**

DeepSeek-R1

updated Jan 21

deepseek-ai/DeepSeek-R1

Text Generation • Updated 12 days ago • 4.16M • 11k

deepseek-ai/DeepSeek-R1-Zero

Text Generation • Updated 12 days ago • 12.9k • 858

deepseek-ai/DeepSeek-R1-Distill-Llama-70B

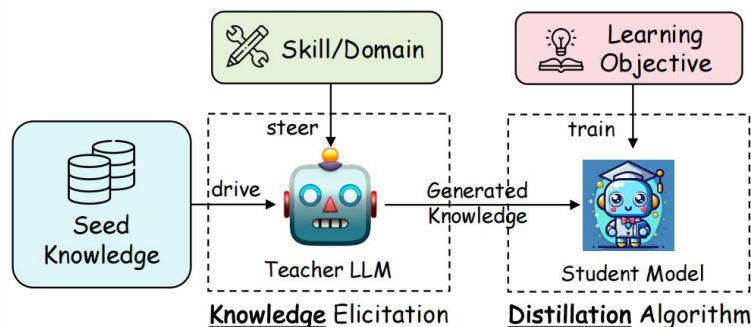
Text Generation • Updated 12 days ago • 435k • 618

deepseek-ai/DeepSeek-R1-Distill-Qwen-32B

Text Generation • Updated 12 days ago • 1.47M • 1.24k

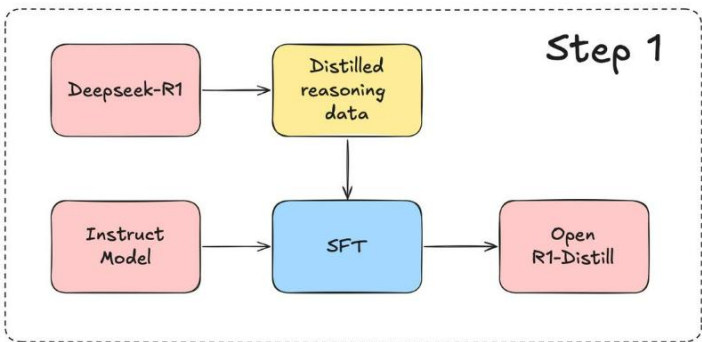
deepseek-ai/DeepSeek-R1-Distill-Qwen-14B

Text Generation • Updated 12 days ago • 636k • 455



● 应用实例- Deepseek的蒸馏与复现

- Deepseek-R1-Distill的蒸馏数据**不是完全开源的**
- Open-R1: 知识获取得到**类官方蒸馏数据**，训练得到近似效果



Model	MATH-500 (HF lighteval)	MATH-500 (DeepSeek Reported)
DeepSeek-R1-Distill-Qwen-1.5B	81.6	83.9
DeepSeek-R1-Distill-Qwen-7B	91.8	92.8
DeepSeek-R1-Distill-Qwen-14B	94.2	93.9
DeepSeek-R1-Distill-Qwen-32B	95.0	94.3
DeepSeek-R1-Distill-Llama-8B	85.8	89.1
DeepSeek-R1-Distill-Llama-70B	93.4	94.5

@51GTO博客

可以使用已开源的类官方蒸馏数据，或者自建数据集，以强化现有学生模型学习教师模型(Deepseek, GPT) 的知识！

Report - [A Survey on Knowledge Distillation of Large Language Models](#)

Report - [Knowledge Distillation: A Survey](#)

知乎 - Deepseek v3 技术报告万字硬核解读
<https://zhuanlan.zhihu.com/p/16323685381>

网页 - [大白话说清楚DeepSeek的蒸馏技术到底是什么？](#)
[大规模语言模型知识蒸馏综述](#)
[万字长文详解大模型知识蒸馏指南](#)
[大模型蒸馏（以tulu3和deepseek-R1为主）](#)
[一文汇总 deepseek R1 最新复现进展](#)

谢谢！ 请多提意见！



大连理工大学
信息检索研究室



Information Retrieval Laboratory of DUT



大连理工大学

信息检索研究室

搜人搜物搜信息 重情重义重认知

<http://ir.dlut.edu.cn>