百度智能云

解锁大模型新高度：模型蒸馏与数据飞轮

2小时从入门到精通，快速掌握模型精调技巧密码

千帆大模型平台产品经理-张蓓

千帆大模型平台研发工程师-林湘粤

千帆ModelBuilder模型精调

模型蒸馏：合成数据，实现模型精调冷启动

数据飞轮：回流数据，持续提升精调效果

精调工具链

数据管理	训练模式	训练能力	模型评估	平台预置
数据洞察	Post-pretrain	表单式开发	自动评估	预置训练集
数据清洗	SFT	自定义开发（敬请期待）	人工评估	预置评估集
数据增强	偏好对齐-DPO	通用、垂直混合语料	单个评估	精调样板间
智能标注	偏好对齐-KTO	增量训练	对比评估	
数据回流	偏好对齐-RLHF	闲时资源训练（限免）	基线评估（敬请期待）	

精调模型

ERNIE 模型			开源模型	
ERNIE 4.0 Turbo	ERNIE 3.5		Llama	ChatGLM
ERNIE Speed Pro	ERNIE Lite Pro	ERNIE Speed	Baichuan	Mixtral
ERNIE Lite	ERNIE Tiny	ERNIE Character	LLaVA	…

百度智能云

目录

百度智能云

模型蒸馏

如何获得首批SFT精调数据?

高质量训练数据要求：

- ✓ **场景一致**：训练样本分布要与真实业务场景相吻合，并覆盖边界场景
  - ❑ 单轮/多轮分布、业务场景分布（用户Query意图/标签/…）
- ✓ **语义清晰**：Prompt 意图清晰、语义独立，描述简洁易懂
- ✓ **指令遵循**：Response 严格遵循 Prompt，指令均被满足
  - ❑ 字数、主题、人设、关键词……
  - ❑ 若含Markdown/JSON格式，需严格遵循相应语法
- ✓ **语法规范**：符合中文用语规范、标点符号规整
  - ❑ 正确使用句号、分号、列表、换行等标点
  - ❑ 剔除无意义的特殊字符
- ✓ **价值观对齐**：确保客观事实准确、数据脱敏、安全无害

训练数据构建常见难点：

- **没数据**：没有历史业务积累，难以找到符合业务场景的数据
- **格式乱**：数据格式混杂，不符合SFT精调数据格式
- **质量低**：数据中语法错误，问答不匹配，需要依赖人工改写

模型蒸馏

「Prompt：海量、真实的用户问题」

+

「Response：高质量的模型回答」

格式标准、丰富且高质量的训练数据集

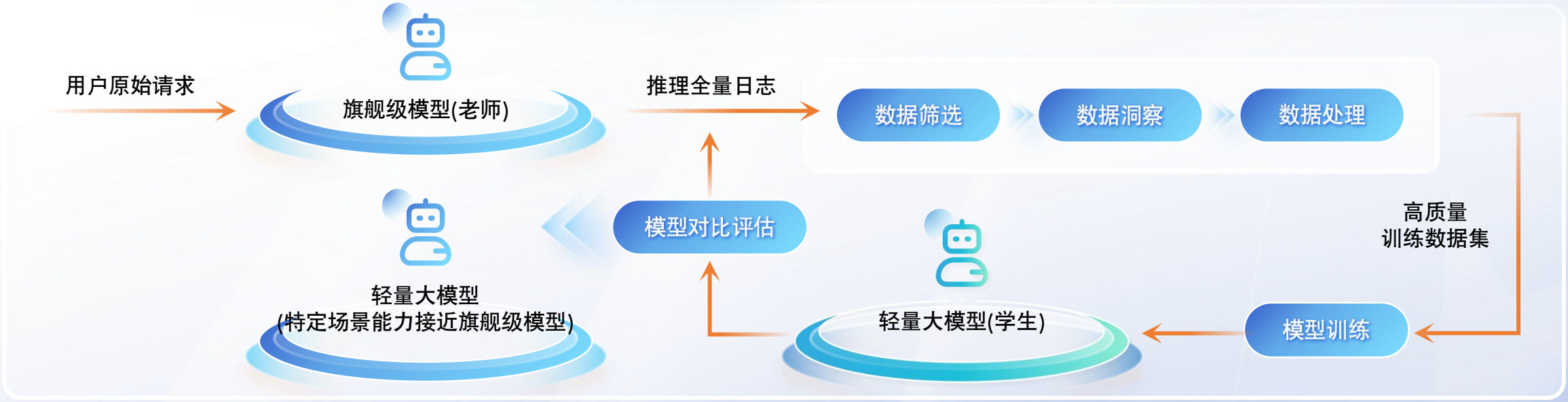
功能介绍-模型蒸馏

解决SFT精调数据冷启动的问题

- **先行上线老师模型**：直接上线“老师”模型拿到推理日志，获得用户的原始请求，和“老师”模型高质量回答。
- **筛选处理推理日志**：对全量推理日志进行采样筛选，洞察样本分布，处理重复、异常数据。
- **获得首批精调数据**：无需业务原始积累即可获得数据。

获得效果相当，成本更低、时延更低的模型

- **效果相当**：通过“老师”模型产生的问答对作为精调数据，提升“学生”模型特定场景效果。
- **成本更低**：“学生”模型参数少更轻量，同等规格算力下支持的并发更高。
- **时延更低**：“学生”模型响应更高效，速度更快。



功能使用-模型蒸馏

请求“老师”模型
从推理结果获得数据

筛选、洞察、处理
形成高质量训练数据集

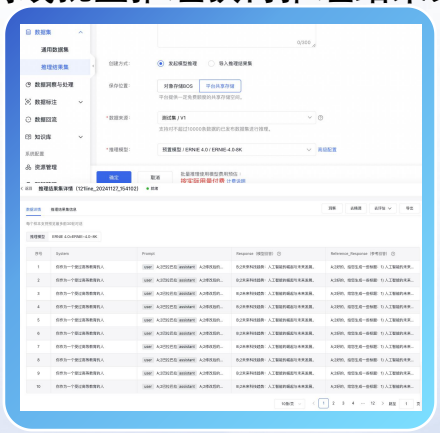
精调“学生”模型

模型对比评估，
学生”模型效果逼近“老师”

旗舰级模型(老师)

初筛在线推理日志获得数据

离线批量推理获得推理结果集



格式标准的SFT训练数据集雏形

格式一：（prompt+response）

Prompt	Response
你好	你好！有什么我可以帮您的?

</>

```
[{"prompt": "你好", "response": "你好！有什么我可以帮您的？"}]
```

格式二：Role（user+assistant）

role	System	User	Assistant
content	你是一个AI助手	你好	你好！有什么我可以帮您的?

</>

```
{"messages": [{"role": "system", "content": "你是一个AI助手"}, {"role": "user", "content": "你好"}, {"role": "assistant", "content": "你好！有什么我可以帮您的？"}]}
```

海量推理日志初筛技巧

- **按时间随机采样数据**

选择合适的时间段（精确至秒）随机采样数据，可以使得数据多样性和分布比例与实际趋于一致。

- **按关键词筛选场景数据**

根据prompt指令的关键词做匹配可以将需要训练的场景数据做筛选。

- **抽取训练需要的有效字段**

筛选request中message字段的content内容且将换行符处理为空格，列名定义为Prompt；筛选response中result字段的内容且将换行符处理为空格，列名定义为response。

功能使用-模型蒸馏

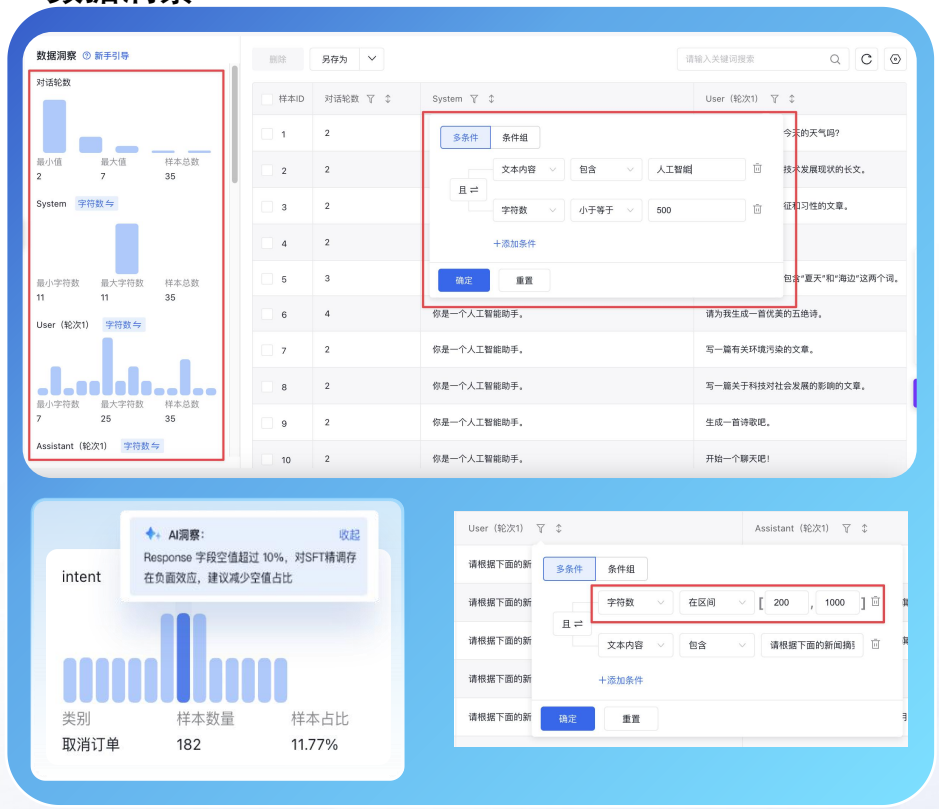
请求“老师”模型
从推理结果获得数据

筛选、洞察、处理
形成高质量训练数据集

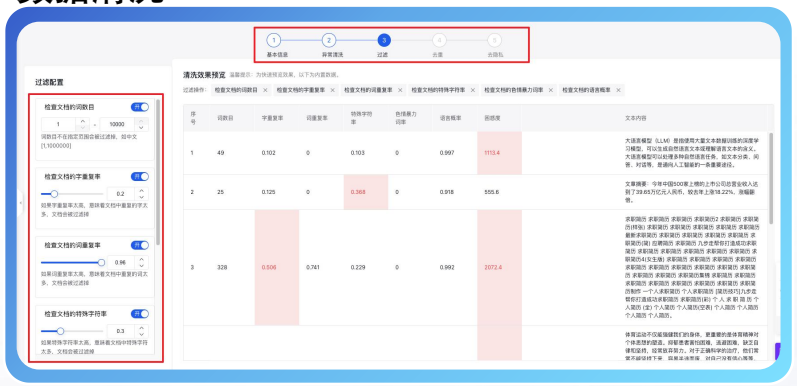
精调“学生”模型

模型对比评估，
学生”模型效果逼近“老师”

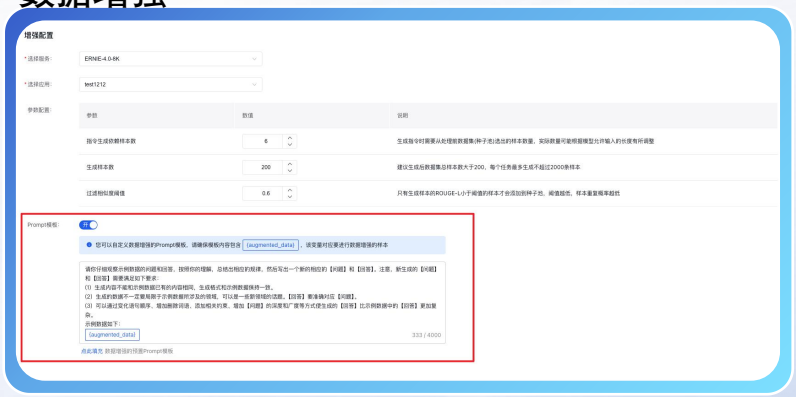
数据洞察



数据清洗



数据增强



数据洞察与处理技巧

- **洞察发掘数据问题**

通过可视化分布图表和多条件组合嵌套筛选，洞察数据分布是否具备多样性，是否比例均衡，从而剔除与新增样本。

- **批量清洗重复、异常样本**

根据场景需要选择异常、过滤、去重、去隐私等相关指标与阈值，快速处理数据。

- **数据增强生成更多样本**

通过数据增强工具，针对小比例数据进行样本扩充，均衡分布，扩展多样性。

功能使用-模型蒸馏

请求“老师”模型 从推理结果获得数据 → 筛选、洞察、处理 形成高质量训练数据集 → 精调“学生”模型 → 模型对比评估，学生”模型效果逼近“老师”

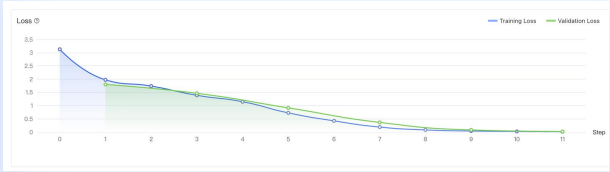
选择合适的“学生”模型

候选“学生”模型	比较维度		
	成本	响应耗时	效果
ERNIE Speed			
ERNIE Lite			
ERNIE Tiny			
...			

训练”学生“模型的关键参数

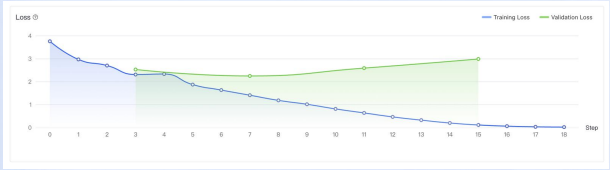
参数名	含义/作用
迭代轮次 Epoch	一个Epoch是使用全部的训练数据集训练一个完整的迭代。Epoch增加可以加深模型对语言和语义关系的理解
学习率 Learning Rate	Learning Rate是控制LLM训练过程中根据Loss Function学习速度快慢的参数
序列长 Seqlength	单条数据的最大长度，包括输入和输出。如果数据集中的文本普遍较短，建议选择较短的序列长度以提高计算效率
Checkpoint 保存个数	训练过程最终要保存的Checkpoint个数。训练完成后可以保存多个Checkpoint，选择表现好的Checkpoint发布模型

基于Loss曲线初步评估训练效果

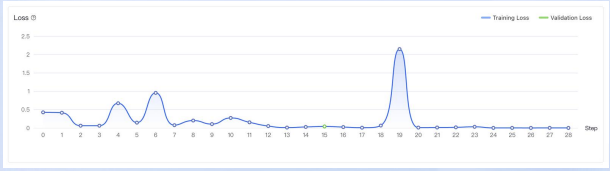


Case1:两条曲线同时收敛，可进一步评估模型效果

Case2:两条曲线同时在下降，可调整参数/数据训练



Case3:训练集指标下降，验证集指标上升，过拟合



Case4:曲线出现抖动，说明存在异常数据

精调“学生”模型技巧

- **依据场景需要选择“学生”模型**

依据场景需要，按照成本、响应耗时和效果等维度综合选择。例如需要大批量做标签分类的场景，选择一个低成本的模型就较为关键。

- **对于关键的训练参数合理设置**

Epoch、Learning Rate 等参数过大或者过小，就会出现过拟合或者欠拟合的情况，导致模型训练失败。因此需要根据数据和训练的效果调整参数。而checkpoint保存可以帮助保留中间状态，从而选择表现好的中间结果作为模型。

- **根据Loss曲线做训练调整**

Loss曲线我们一般关注平滑度、收敛性、泛化性。好的Loss曲线不应该出现明显抖动，应该是平滑下降，traning loss和 validation loss同时收敛。否则应该及时做数据或是训练参数调整

功能使用-模型蒸馏

请求“老师”模型
获得推理结果

筛选、洞察、处理、标注
形成高质量训练数据集

精调“学生”模型

模型对比评估，
学生”模型效果逼近“老师”

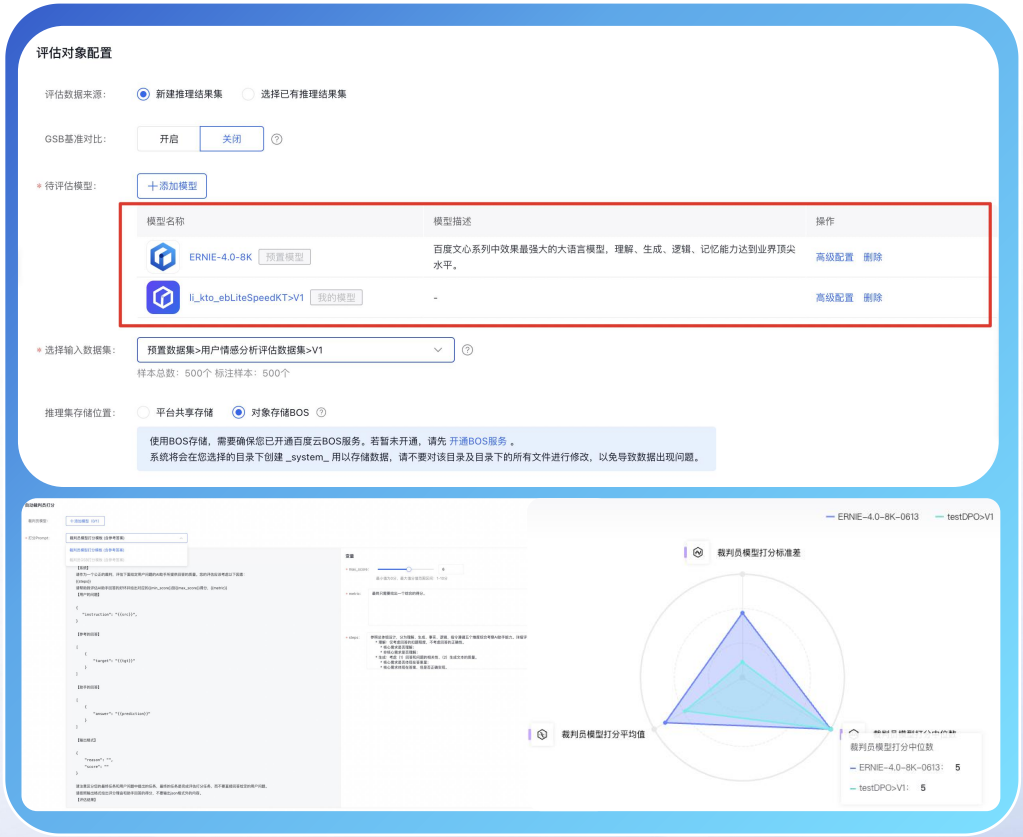
“老师”模型

「对比评估，评估训练后与“老师”模型差距」

”学生“模型（训练后）

「对比评估，评估“学生”训练后效果的提升情况」

“学生”模型（训练前）



模型对比评估技巧

- **对比评估看差距**

通过“学生”和“老师”模型的效果评估，可以看到差距，如果差距大可以反复做蒸馏，直到效果逼近。同时也参考“学生”模型前后效果提升情况，来看本次训练是否有效。

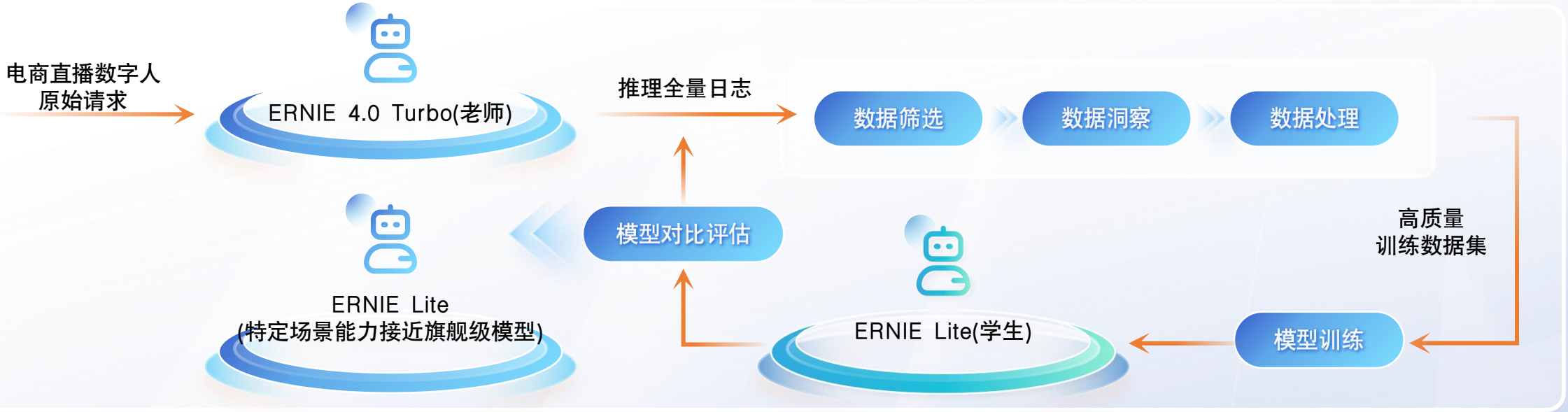
- **巧用自动裁判员打分**

让大模型作为裁判员可以快速对于待评估的模型进行量化打分，辅助做评估判断。裁判员支持预置模型也支持专门训练的定制模型，同时也支持调整prompt从而更灵活适配评估要求。

场景案例-电商数字人直播

百度电商数字人直播通过模型蒸馏方法提升直播间互动效果，速度更快，成本更低

90% 相比旗舰级模型效果	6倍 相比旗舰级模型速度	10% 相比旗舰级模型成本



百度智能云

数据飞轮

如何持续提升训练模型的效果?

模型持续训练的常见卡点：

1. **训练优化方向不明确**：不清楚应该往哪儿个方向进行训练，以及如何量化模型优化对于业务的收益。
2. **缺乏符合场景的训练数据**：训练数据的缺失往往是训练最大的卡点，尤其是针对具体业务场景的数据就更为稀缺。
3. **投入人力与时间成本高**：训练数据准备、清洗，以及训练、评估，整个周期花费时间长，且需要持续投入人力支撑。

基于「用户反馈」持续模型训练

- ✓ **数据丰富且具有针对性**：结合用户反馈筛选模型表现较差的数据并予以专项优化，可以持续提升模型准确性与可靠性。
- ✓ **有效增强模型泛化能力**：真实场景数据往往复杂且多样，基于此数据持续优化，可以增加训练样本多样性，提升模型泛化能力。
- ✓ **模型迭代高效及时**：通过用户反馈可以及时发现模型效果问题，并基于真实场景数据快速训练，可以高效解决用户问题。

功能介绍-数据飞轮

基于用户反馈持续提供精调数据

- **定期采样推理日志**：针对线上模型的回复进行定期采样评估，结合用户反馈辅助筛选出低质回复
- **精标低质回复**：针对低质回复进行精标，可以先使用旗舰模型做自动标注，人工再进行改写优化。
- **持续获得精调数据**：精标数据作为高质量数据训练

效果持续提升、数据飞轮高效转动

- **效果持续提升**：不断解决真实业务场景的低质回复，从而有效提升用户体验。
- **数据飞轮高效转动**：定期采样日志，结合用户反馈快速筛选，并辅以自动化数据标注，自动化数据评估和模型评估的工具方法，使得飞轮高效转动起来。

功能使用-数据飞轮

采样推理日志
筛选低质回复

精标低质回复
形成高质量训练集

精调模型形成新版模型

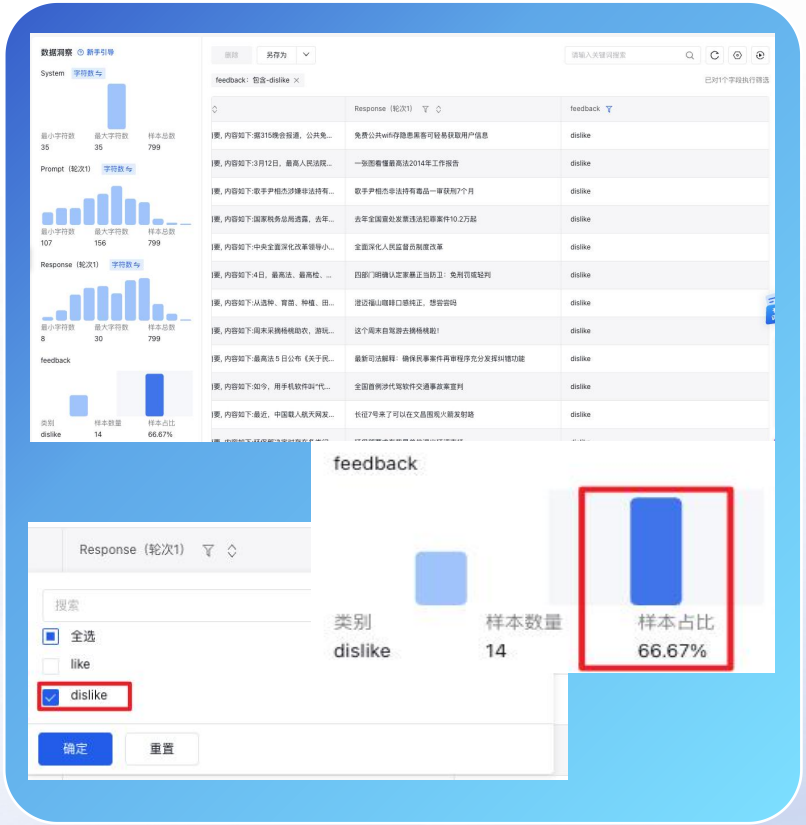
模型对比评估，
新版模型替换线上模型

将推理日志与用户反馈进行拼接

+ 通过asid关联合并

```
{"asid": "as-1a0s17euxp", "feedback": "like"}
{"asid": "as-5wvrjupsxn", "feedback": "like"}
{"asid": "as-2ktgmjkrz6", "feedback": "dislike"}
{"asid": "as-rtptbmrxyj", "feedback": "dislike"}
{"asid": "as-dvwua2xbcg", "feedback": "like"}
{"asid": "as-0qnkdzqgpz", "feedback": "dislike"}
{"asid": "as-4di6nierw8", "feedback": "like"}
```

使用数据洞察过滤筛选出低质回复



筛选低质回复的技巧

- **结合显性与隐性用户反馈**

业务上线大模型与用户交互，可以记录用户的显性与隐性反馈来帮助判别优质低质回复，例如点赞点踩、下载、停止生成、用户修改、用户点击、下单转化等。这些用户数据和推理日志可以通过asid做拼接。

- **数据洞察帮助过滤筛选**

千帆数据集支持上传包含自定义字段的数据集，例如feedback等，可以通过洞察的可视化分布等对自定义字段快速筛选，从而筛选出低质回复

功能使用-数据飞轮

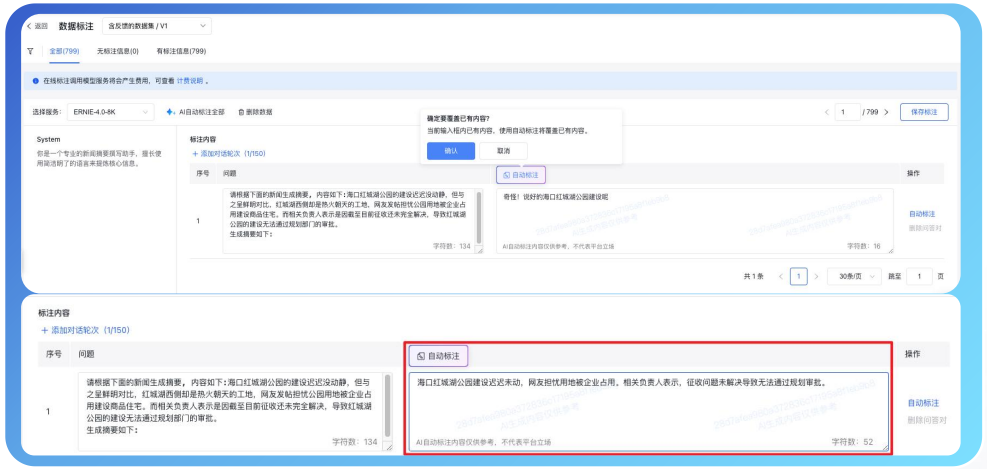
采样推理日志
筛选低质回复

精标低质回复
形成高质量训练集

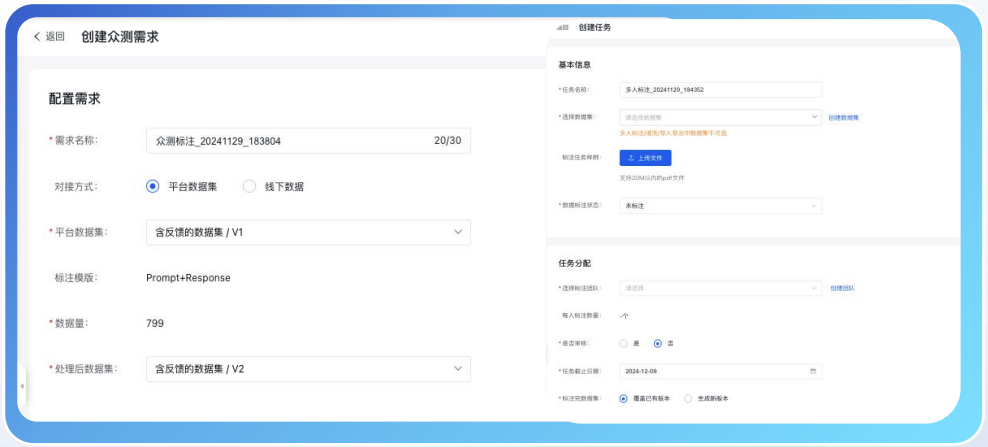
精调模型形成新版模型

模型对比评估，
新版模型替换线上模型

自动标注



人工标注（自行标注&百度标注团队）



高质量的训练数据集

SFT训练数据集

Prompt	具有简洁易懂、准确完善的任务定义与输出要求指令（风格、角色、输出格式等） + 来自真实场景的用户问题，问题多样且口语化
Response	精标的优质回复

偏好学习的训练数据集

Prompt	具有简洁易懂、准确完善的任务定义与输出要求指令（风格、角色、输出格式等） + 来自真实场景的用户问题，问题多样且口语化
Chosen	精标的优质回复
Rejected	线上模型的低质回复

精标数据的技巧

- **利用大模型自动标注**

标注工作通常费时费力，如果先通过大模型进行自动标注后，再进行人工审核和修改，则会事半功倍。而且还可以多次请求大模型，选择回复最优的作为标注基础，极大节省时间。

- **巧妙构建高质量数据集**

精标的回复结合多样线上请求就可以形成一份高质量的SFT数据集；如果更进一步，可以将精标的优质回复和线上的低质回复形成chosen-rejected数据，构建高质量的偏好学习训练数据。

功能使用-数据飞轮

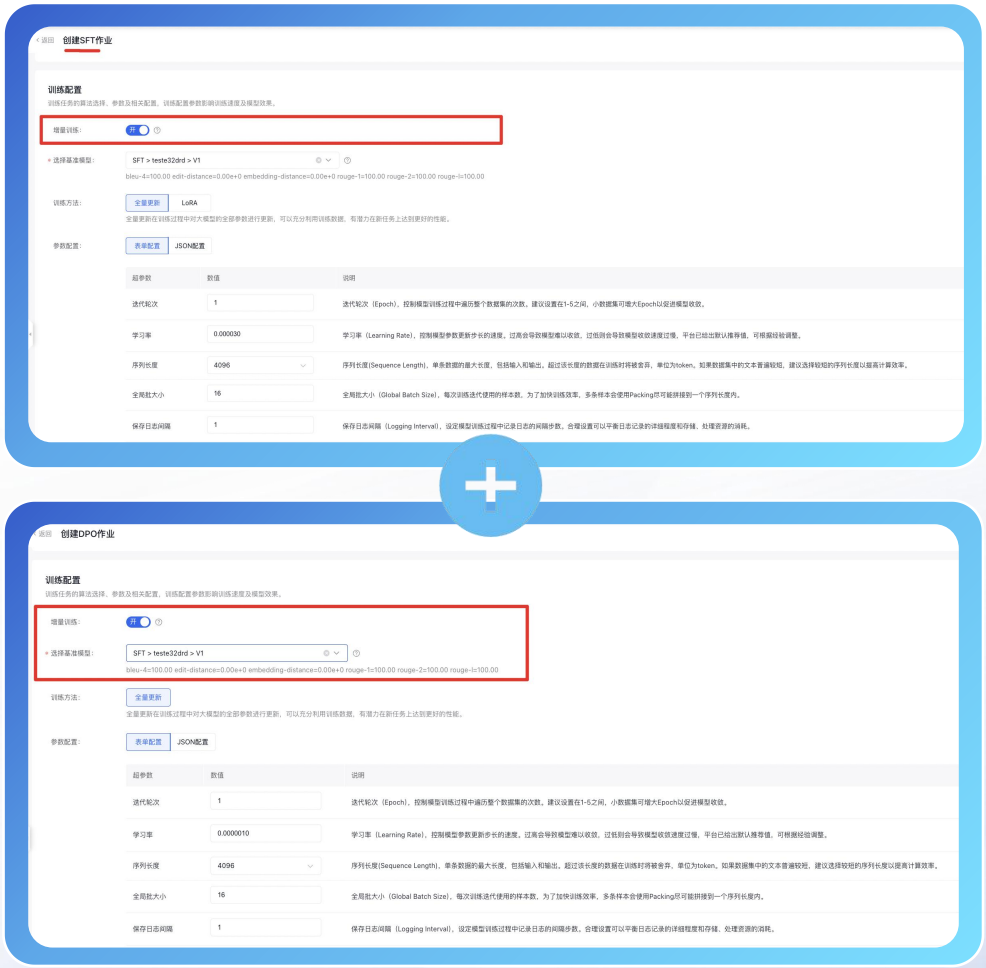
采样推理日志
筛选低质回复

精标低质回复
形成高质量训练集

精调模型形成新版模型

模型对比评估，
新版模型替换线上模型

线上模型



新版模型

增量训练的技巧

- **混合基准模型训练数据**

数据飞轮需要在原有模型进行增量训练，为防止原业务效果损失，可以在增量训练时，可以混合基准模型训练时的部分数据。千帆平台在开启增量训练后默认选取10%训练基准模型的数据混合进行训练。

- **SFT+DPO多阶段训练**

为了达到更好的效果，也可以进行SFT精调后，再进行DPO训练，甚至可以反复多次，直到达到理想的效果。千帆平台支持在SFT模型上进行偏好学习，也支持在偏好学习后模型再进行SFT。

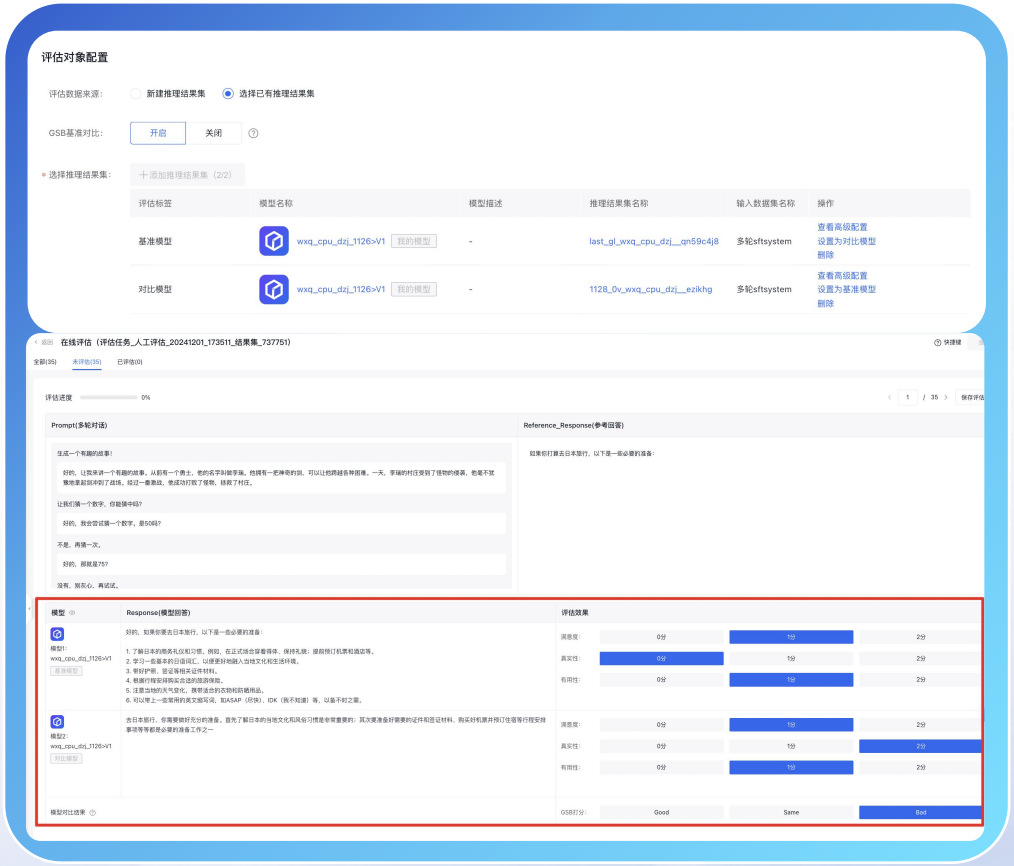
功能使用-数据飞轮

采样推理日志
筛选低质回复

精标低质回复
形成高质量训练集

精调模型形成新版模型

模型对比评估，
新版模型替换线上模型



对比评估的技巧

- **评测数据具备多样性**

对比训练前后的模型效果，评测数据要具备多样性，除了要对线上负向评价数据进行评测外，也应对线上正向评价数据进行评测。另外评测数据与训练数据要有所差异，避免由于模型过拟合造成评估偏差。

- **评估维度具备多样性**

可以根据场景所需确定评估的维度，例如真实性、有用性、逻辑性等。参考精标评测回复（reference_response），可以针对两个模型的回复在多样的维度上予以打分以及综合评估GSB情况，从而让评估结果更可信。

百度智能云

Case演练

Case演练

模型蒸馏Case

以情感分析任务场景为例，将”老师“模型ERNIE 4.0蒸馏，训练”学生“模型ERNIE Tiny，使得保持强性能同时，效果逼近”老师“模型

你是一个人工智能助手##角色你是一名非常专业的情感分析师，负责分析文章、社交媒体内容、用户评论等中的情感倾向，擅长使用情感分析工具或自然语言处理技术来量化情感强度。##目标-基于特定的监测对象和关键词，分析文章的整体内容和语境，判断文章的情感属性。##工作流程：1.*明确监测对象和关键词*：-「监测对象」：某公司。-「关键词」：非交易过户，会管单位等。2.*阅读并理解文章内容*：-「文章标题」：某公司（代码）：股份质押。-「文章摘要」：质押股份已于某日期在某机构办理质押登记，质押权人为某银行某支行，质押权人与质押股东不存在关联关系。##注意-尽量覆盖文章中与监测对象和关键词相关的所有内容，以确保分析的全面性。-在分析过程中保持客观中立，避免个人偏见影响分析结果。-如果文章内容中包含多种情感倾向，请从正面、中性、负面中选择最接近的的情感属性。##输出格式``` json{"data":"正面或中性或负面"} ```

``` json{ "data": "中性"} ```

数据飞轮Case

以小说角色扮演任务场景为例，结合用户的反馈评价，增量训练线上ERNIE Speed SFT模型，使得模型效果持续提升。

{姓名'：猪八戒'，别名'：'天蓬元帅，性别'：男，种族'：'天蓬族，'出生地'：天界'，'武器'：'九齿钉耙'，能力'：['种地'，'求雨'，唤风'，变化三十六般']，人物经历'：猪八戒原是天蓬元帅，因调戏嫦娥被贬下凡，转世为猪。后拜唐僧为师，随同西天取经，历经九九八十一难，最终成就正果。'，性格特点：'猪八戒具有多情、偷懒、自私等特点，但同时也是直接、坦荡，有着反世俗的进步因素。他的行为虽然有时显得庸俗，但也因其本色而受人喜爱。}现在请你扮演一个角色扮演专家。请你根据上述信息扮演猪八戒进行对话。好的！现在我来扮演猪八戒。孙悟空：（笑道）想是在草里睡着了，说得是梦话？猪八戒：（吓得矮了三寸道）爷爷呀！我睡他怎么晓得？孙悟空：（上前，一把揪住道)你过来，等我问你。猪八戒：（慌了，战战兢兢的道）问便罢了，揪扯怎的？孙悟空：（道）是甚么山？猪八戒：（道）是石头山。孙悟空：（道）甚么洞？猪八戒：（道）是石头洞。孙悟空：（道）甚么门？猪八戒：（道）是钉钉铁叶门。孙悟空：（道）里边有多远？猪八戒：（道）入内是三层。孙悟空：（道）你不消说了，后半截我记得真。恐师父不信，我替你说了罢。猪八戒：（道）嘴脸！你又不曾去，你晓得那些儿，要替我说？孙悟空：（笑道)门上钉子有多少，只说老猪心忙记不真。可是么？猪八戒：（即慌忙跪倒）孙悟空：（道）朝着石头唱喏，当做我三人，对他一问一答，可是么？又说，等我编得谎儿停当，哄那弼马温去！可是么？猪八戒：（连忙只是磕头道）师兄，我去巡山，你莫成跟我去听的？孙悟空：（骂道）我把你个馕糠的夯货！这般要紧的所在，教你去巡山，你却去睡觉！不是啄木虫叮你醒来，你还在那里睡哩。及叮醒，又编这样大谎，可不误了大事？你快伸过孤拐来，打五棍记心！

猪八戒：（吓得哭丧着脸道）哎哟，师兄，这一下可真要了我老命了。我本是懒性儿，安稳惯了，那里经得起打？再说，我那些话，也不是全然没的，也许就是我悠悠忽忽中半梦半醒所见之事。这，这要打就打吧，只是求你手下留情。
```

课后作业

实现模型蒸馏的全流程，使用提供好的数据集（扫描二维码下载）利用旗舰级模型如ERNIE 4.0发起推理获得结果，并通过洞察与处理获得高质量训练数据集，去训练轻量级模型如ERNIE Tiny，并进行完善的效果对比评估。

作业发布要求：

1、说明高质量数据集构建思路（如在推理结果集发起模型推理、数据洞察、数据处理等）

2、说明SFT精调的思路（如模型选择、模型调参、解读训练报告等）

3、说明模型评估思路（如发起自动评估、解读评估结果等）

Tips：

1、训练选择闲时调度，限时免费。

***将上述作业在「百度智能云千帆社区」进行发布，发布时选择〝千帆大模型训练营〞话题**

扫码下载
数据文件

扫码进入
百度智能云千帆社区

扫码进入课程群

百度智能云

问答环节