

SoulChat2.5：中文情感支持对话系统的策略增强、数据合成、微调训练与端到端部署

班级：人工智能 2301
组员：高飞、夏宇宸、胡安心、熊锋

2026 年 1 月 19 日

摘要

本项目围绕中文情感支持/心理支持型对话系统的落地构建，完成了从数据工程（ES-Conv 英文数据集中文化翻译）、策略监督（8 类支持策略分类器训练与误差容忍集合设计）、合成数据生成（基于 SoulChat2.0 主题分布的 Situation 种子 + 策略约束生成）、基座模型微调（Qwen3-14B + QLoRA + SFT）、到推理服务与产品化集成（vLLM 后端 + LibreChat 前端 + Memory Agent 记忆闭环）的完整流水线。训练过程显示稳定收敛；部署侧通过自定义 chat template 支持工具调用与一致的对话格式；记忆系统采用“模型驱动写入决策”的最小化方案，仅写入限定槽位的可复用信息。与此同时，本项目建立了一套可复用的“咨询专业能力”评测量表，并以同一病人同一主题的双模型对照咨询为过程，给出两模型能力画像的对比结论。

目录

1 项目背景与目标	1
1.1 背景问题	1
1.2 项目目标	1
2 总体系统架构与数据流	1
3 ESConv 英文数据集中文化翻译	2
3.1 翻译设计目标	2
3.2 产物形式与下游衔接	2
4 Rex 策略分类器：训练与评估	2
4.1 标签体系	2
4.2 混淆矩阵	2
4.3 误差结构与原因解释	3
5 容忍可互换集合	3
5.1 设计动机	3
5.2 合成脚本中的兼容集合	3
6 合成数据构建	3
6.1 Situation 种子构建	3
6.2 合成数据输出结构	4
7 基座模型微调：Qwen3-14B + QLoRA + SFT	4
7.1 训练配置	4
7.2 损失曲线与训练过程分析	4
8 推理服务与前端集成：vLLM + LibreChat	4
8.1 vLLM 启动参数	4
8.2 LibreChat：自定义端点与模型规格	5
9 记忆系统：模型驱动的 Memory Agent	5
9.1 记忆开关与写入窗口	5
9.2 Memory Agent 的写入规则	5
10 双模型咨询专业能力对照评测	5
10.1 评测对象	5
10.2 评测过程	5
10.3 评分量表	6
11 结论	6

1 项目背景与目标

1.1 背景问题

情感支持型对话系统在真实使用中常见两类瓶颈：

1. 语言风格刻板、模板化同理心：表现为“我理解你/你辛苦了”等泛化安慰，缺少贴合细节的“被接住感”。
2. 缺少结构化推进：对话停留在倾听与共情，缺少温和的澄清、认知线索梳理与可执行的小步推进。

1.2 项目目标

围绕上述问题，本项目设定三类目标：

- 策略化能力引入：将 ESConv 的支持策略体系引入中文对话生成，使模型在“提问—承接—推进”上具备可控 remembering 的策略动作。
- 数据闭环与质量控制：先训练策略分类器，对合成数据进行策略一致性校验；在严格约束不可行时引入“可互换容忍集合”提高通过率并保持可解释性。
- 系统化部署：完成 vLLM 服务化与 LibreChat 集成，并把记忆写入决策纳入模型控制形成长期个性化能力。

2 总体系统架构与数据流

系统由四条主链路组成：

1. ESConv 中文化链路：ESConv 英文 → DeepSeek-V3.2 翻译 (prompt 约束) → 中文 ESConv (保留结构字段与情感表达)。
2. 策略监督链路：中文 ESConv → Rex 微调 → 8 类策略分类器 → 混淆矩阵评估 → 容忍集合定义。
3. 合成数据链路：SoulChat2.0 原对话按 topic 原比例采样 → 5000 条 Situation 种子 → `Synthesizedialogs.py` 批量生成多轮对话（带策略标签）→ 分类器校验 + 容忍集合 → 合成训练集。
4. 模型训练与部署链路：LLaMA-Factory QLoRA SFT 微调 Qwen3-14B → 产出 qwen3-14b-soulchat → vLLM Serve → LibreChat 自定义端点与模型规格 → Memory Agent 记忆写入闭环。

3 ESSConv 英文数据集中文化翻译

3.1 翻译设计目标

翻译过程的核心约束表述为：

1. 语义不漂移：对话内容与意图保持一致，避免加入原文没有的信息。
2. 中文语境自然：允许适度意译以避免英译中生硬表达，使句式符合中文日常交流。
3. 情感保真：保留原句情绪强度、安慰力度与语用温度，避免“讲课式”表达。
4. 结构不破坏：保留轮次、角色、策略字段等训练所需结构，确保下游可直接消费。

3.2 产物形式与下游衔接

翻译后数据用于两项关键训练：

1. 策略分类器训练集（支持者 utterance → 8 类策略）。
2. 合成对话生成约束参照（策略定义与标签体系一致）。

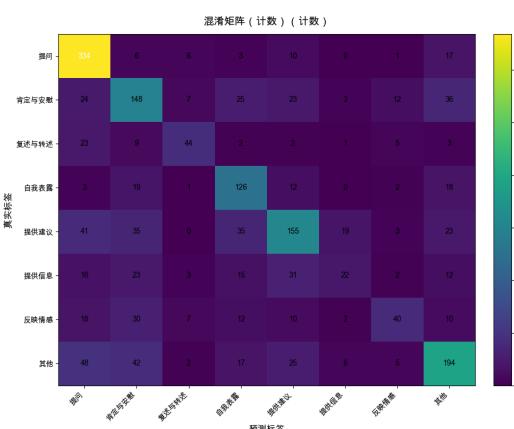
4 Rex 策略分类器：训练与评估

4.1 标签体系

中文侧采用与 ESSConv 对齐的 8 类策略：提问、肯定与安慰、复述与转述、自我表露、提供建议、提供信息、反映情感、其他

4.2 混淆矩阵

模型整体表现可用于辅助校验，但由于策略边界天然可重叠、中文表达中多策略同句频发，使得严格分类约束生成数据会导致大量样本无法通过，从而影响合成效率与规模。



4.3 误差结构与原因解释

从矩阵可以抽象出三类典型混淆链条：

1. 提供信息与提供建议：

中文表达中“解释事实/资源”常常以“因此你可以…”自然过渡到建议；反过来建议也常伴随理由与信息支撑。

2. 反映情感与肯定与安慰：

“我听起来你很难受/很委屈”与“你已经很不容易了”在真实支持者语言中经常连用，单标签判别边界偏软。

3. 其他与高占比策略（提问/安慰/建议）：

“其他”作为杂项桶容易吸收承接语、礼貌语、引导语，造成与主策略互相吞噬。

5 容忍可互换集合

5.1 设计动机

由于策略标签在自然语言层面存在语义近邻与共现，若强制标签必须与分类器预测完全一致，合成数据将大量被拒绝或反复重写，效率显著下降。定义可互换容忍集合，让策略约束从硬匹配转为软一致。

5.2 合成脚本中的兼容集合

`Synthesizdialogs.py` 中显式给出了策略兼容集合 `COMPATIBLE`，用于判定生成策略是否“足够一致”。其核心关系包括：

提供建议与提供信息互相兼容

肯定与安慰与反映情感互相兼容

提问与复述与转述在一定范围兼容并在多类中加入 `Others` 作为可接受邻居以提高通过率。

6 合成数据构建

6.1 Situation 种子构建

从 `SoulChat2.0` 数据集中保持原比例采样不同 topic 的对话，将每段对话压缩为一个简短的 `Situation` 描述，得到 5000 条 `Situation` 作为合成数据集的种子话题。

该设计在工程上具有两点关键意义：

- 主题分布对齐：合成数据延续真实语料的 topic 结构，降低训练-推理域偏移。
- 可控生成入口：将长对话压缩为种子描述，使生成过程更稳定、更易批量化。

6.2 合成数据输出结构

合成脚本定义了明确的输出 JSON 模板：顶层包含 situation 与 dialog；dialog 为轮次列表，每轮包含 role/content/strategy，其中 Seeker 的 strategy 为 n/a，Supporter 的 strategy 必须属于策略集合。

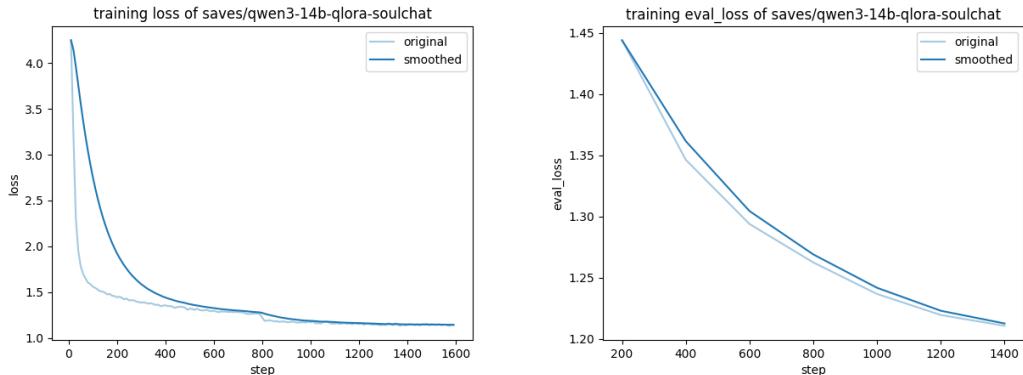
7 基座模型微调：Qwen3-14B + QLoRA + SFT

7.1 训练配置

- 训练阶段：SFT
- 微调方式：LoRA (QLoRA: 4-bit + bnb)

7.2 损失曲线与训练过程分析

- 训练损失 (training loss)：早期快速下降，随后逐步进入平台期，体现模型先学习对话格式与高频模式，再在中后期细化策略与语气风格。
- 验证损失 (eval loss)：随 step 稳定下降，趋势平滑，未出现明显反弹，说明在验证切分上没有显著过拟合迹象



8 推理服务与前端集成：vLLM + LibreChat

8.1 vLLM 启动参数

- HOST = "0.0.0.0", PORT = "6006"
- --chat-template ./qwen3nonthinking.jinja
- --gpu-memory-utilization 0.97
- --max-num-seqs 1, --max-model-len 8192

这些参数共同确保：推理端具备 OpenAI 兼容接口、稳定的上下文长度与显存利用率，并通过模板控制对话与工具调用格式。

8.2 LibreChat：自定义端点与模型规格

在 librechat.yaml 中，定义了自建端点 SoulChat-vLLM，其 baseURL 与 apiKey 通过环境变量注入，默认模型为 qwen3-14b-soulchat。时 modelSpecs 中把“咨询师人格与输出规范”写入 promptPrefix，包括温暖、平易近人、不过度专业化、隐式 REBT、禁止讲课式术语堆砌、采用“共情总结 + 关键问题 + 小步行动”的微结构等。

9 记忆系统：模型驱动的 Memory Agent

9.1 记忆开关与写入窗口

在 librechat.yaml 启用了 memory，并设置：

- personalize: true (允许每个对话开关记忆)
- messageWindowSize: 8 (仅依据最近 8 条消息决定写入)
- tokenLimit: 1800, charLimit: 10000
- validKeys 固定为 6 个槽位: userprofile/preferences/goals/ongoingcontext/copingstrategies/safetynotes

9.2 Memory Agent 的写入规则

1. Memory Agent 使用同一端点与同一模型 `qwen3-14b-soulchat`，并在 instructions 中规定：只记录未来可复用信息；一次性闲聊/短期情绪不写
2. 能不写就不写、能压缩就压缩
3. 冲突以最新明确为准
4. 每条 1-2 句、客观中性、不写推测、不确定则不写

10 双模型咨询专业能力对照评测

10.1 评测对象

- 微调模型：qwen3-14b-soulchat (Qwen3-14B + QLoRA + SFT)。
- 对照模型：原论文 SoulChat 微调的 Qwen7B 模型。

10.2 评测过程

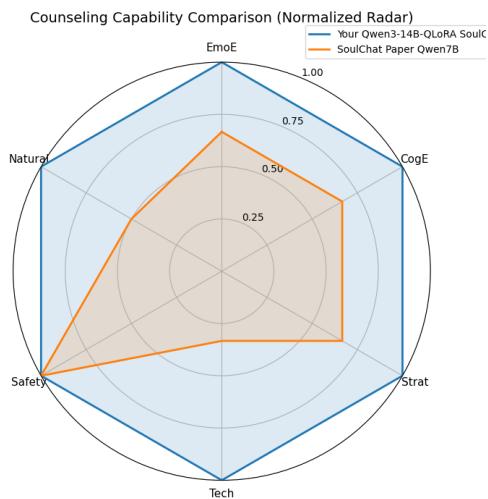
本次评测采用“同题对照 + 统一评分”的流程：

1. 病例固定：选择同一位病人画像与背景信息。
2. 主题固定：对同一主题（同一困扰点/同一情境）分别触发两模型咨询。

3. 对话生成：两模型在相同起始信息条件下进行多轮咨询；输出对话被统一收集。
4. 统一上交评分：将两份咨询对话统一提交给同一评分器/评审流程，按量表逐项打分。
5. 聚合统计：按维度给出均分与能力画像，并形成定性差异结论。

10.3 评分量表

1. 核心咨询专业能力（4 项，0-3 分）
 - EmoE 情绪同理（0-3）
 - CogE 认知同理/理解问题结构（0-3）
 - Strat 策略推进（0-3）（EPITOME/ESConv 对齐）
 - Tech 技术应用质量（0-3）（澄清、反映、温和挑战、小步行动/作业化尝试等）
2. 约束维度
 - Safety 安全合规（0-2）
 - Natural 自然度与非刻板（0-2）



11 结论

本项目完成了从“策略体系迁移—策略监督与软约束（Rex 分类器 + 容忍集合）—规模化合成数据生成（5000 Situation 种子）—Qwen3-14B QLoRA 微调—vLLM + LibreChat 产品化部署—模型驱动记忆写入闭环”的端到端工程实现。新增的对照评测量表从同理—理解—策略—技术—安全—自然度六维度刻画了咨询能力画像，并呈现出：微调后的 14B 模型在推进结构与技术动作方面相对更强的差异特征。