

# DeepSeek-R1 安卓手机部署

“DeepSeek-R1-Distill-Qwen-1.5B 安卓手机部署”

徐静

发布时间：2025 年 02 月 03 日

<https://github.com/DataXujing/DeepSeek-R1-Android>

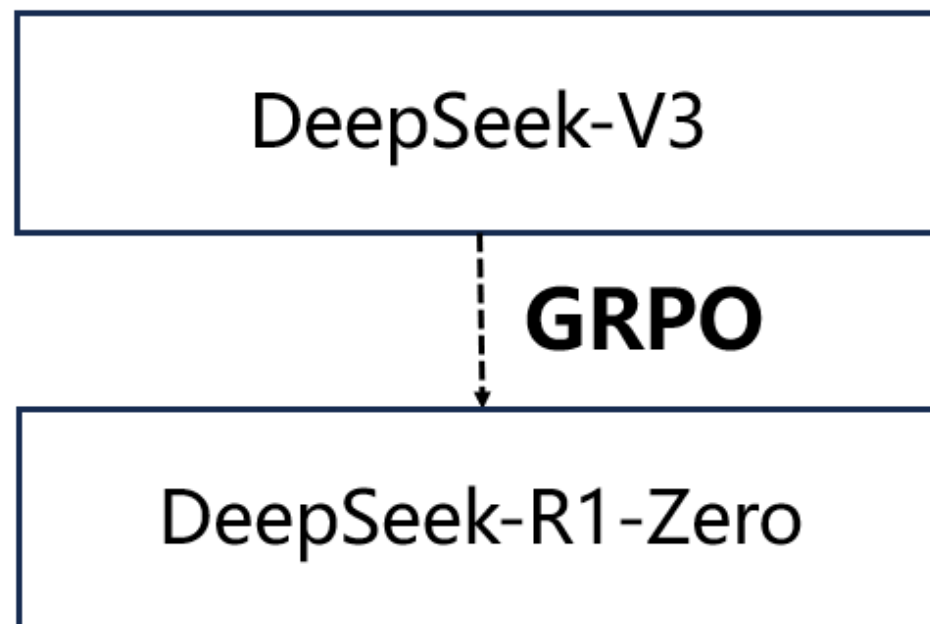
# CONTENT

## Outline

- 1 DeepSeek-R1模型报告解读 [↗](#)
- 2 DeepSeek-R1-Distill-Qwen-1.5B模型下载和转MNN [↗](#)
- 3 基于QT和MNN的DeepSeek-R1安卓端部署 [↗](#)
- 4 Demo [↗](#)

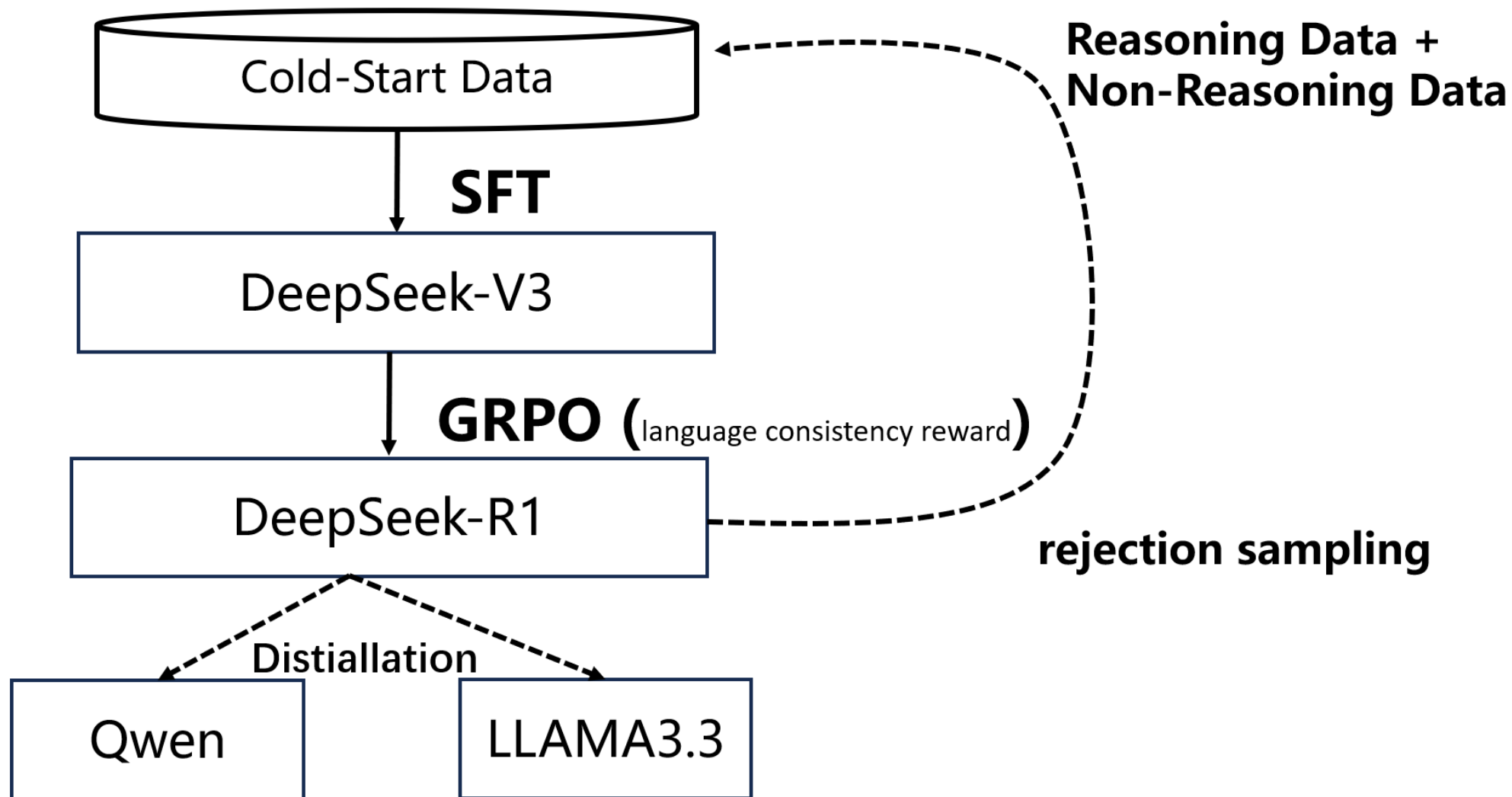
# 1. DeepSeek-R1模型报告解读

# 1.DeepSeek-R1-Zero



- ❑自我进化能力，初步具备了**Test-Time Reasoning**
- ❑Reasoning过程可读性差，中英文混淆
- ❑奖励函数：Accuracy Reward, Format Reward

## 2.DeepSeek-R1



## 3.蒸馏DeepSeek-R1

### 一、确定模型角色

- **教师模型**：选择性能卓越、知识丰富的DeepSeek-R1作为指导者。
- **学生模型**：选用参数适中、易于蒸馏的Qwen-32B作为学习者。

### 二、准备数据集

- **数据收集**：整合涵盖广泛领域的高质量文本数据，为蒸馏过程奠定基础。
- **数据预处理**：严格筛选和处理数据集，确保蒸馏效果不受数据质量影响。

### 三、模型训练与蒸馏

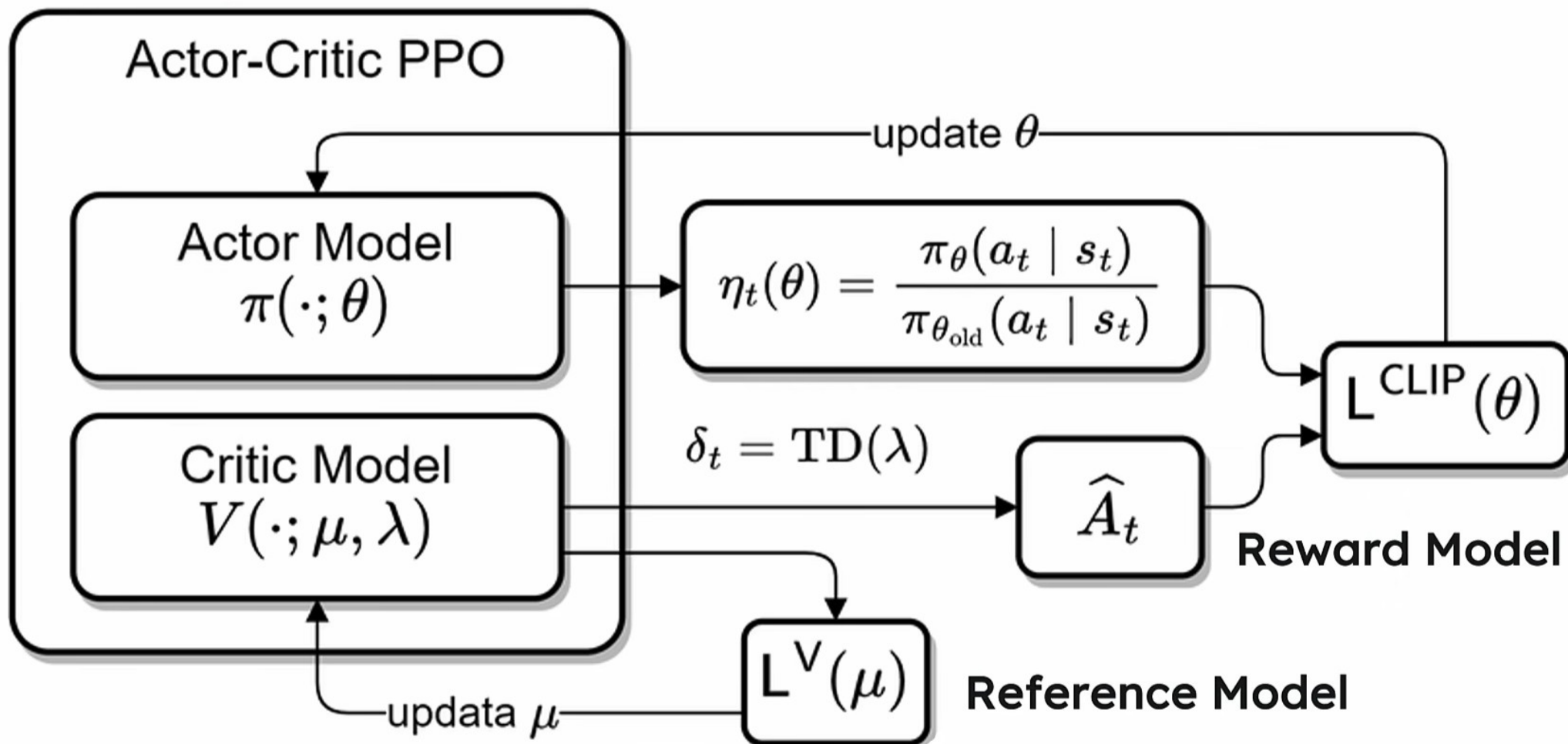
1. **教师模型推理**：DeepSeek-R1输出软标签和中间层特征，作为Qwen-32B的学习标杆。
2. **学生模型训练**：Qwen-32B模仿DeepSeek-R1的决策，继承其知识和经验。
3. **损失函数设计**：采用KL散度或交叉熵等，衡量学生模型与教师模型的输出差异。
4. **优化算法选择**：选用SGD或Adam等算法，更新学生模型参数，最小化损失函数。

### 四、模型评估与调优

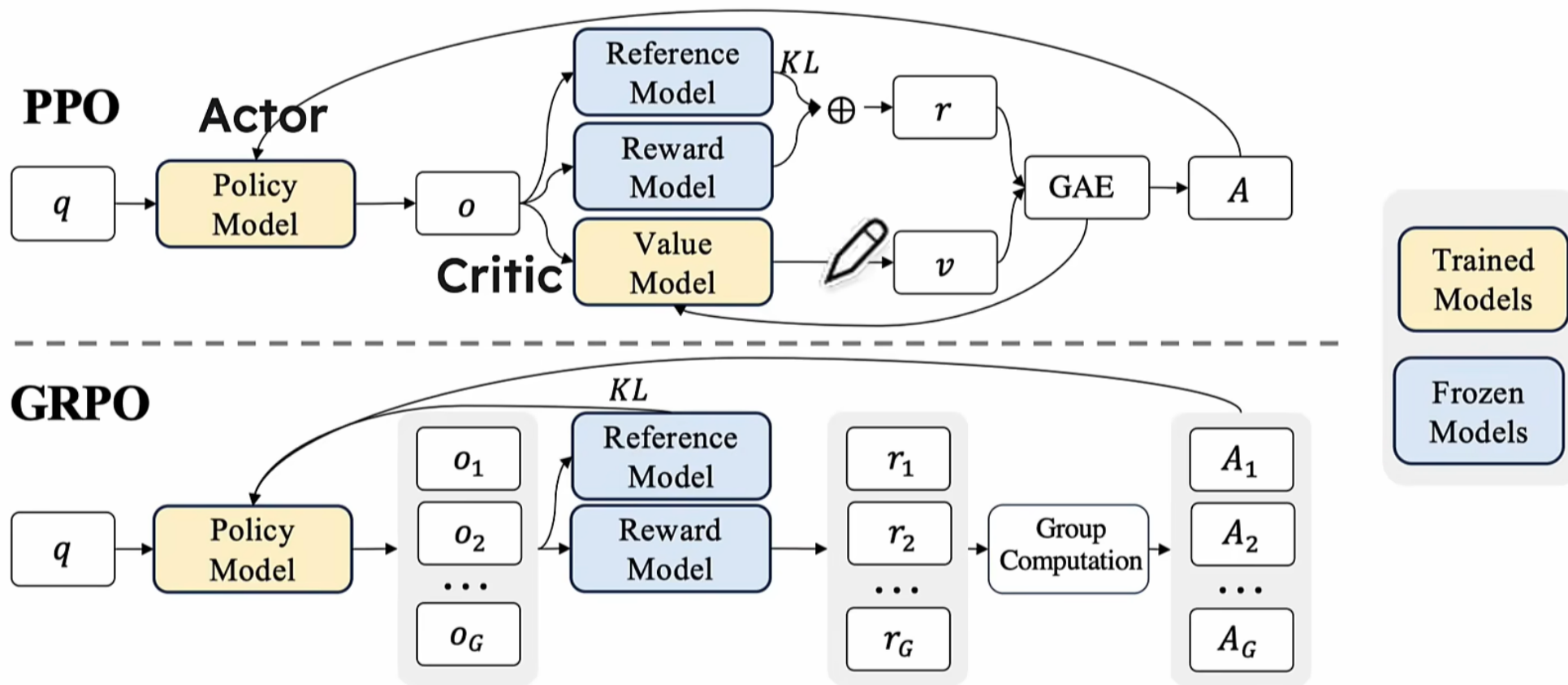
- **模型评估**：利用独立测试数据集，比较学生模型与教师模型的表现，评估蒸馏效果。
- **模型调优**：根据评估结果，调整超参数、优化损失函数等，提升蒸馏效果。

## 4.GRPO

PPO:



# Group Relative Policy Optimization (GRPO)





# RL + LLM 知识:On Policy 与 Off Policy 差异在哪?

- **on policy (PPO):**

- 每次训练都基于自己的生成模型(Actor), 通过教练(Critic)反馈奖励优势
- 优势是效率高, 没有模型自生成自然效率高, 问题是训练后模型能力可能不够

- **off policy (DPO):**

- 基于现有标注的情况进行分析, 存在训练样本可能与模型不匹配的问题
- 优势是更有可能达到模型能力的上限, 问题是效率较低

# PPO存在的问题

从PPO的优化过程分析，其存在如下缺点：

1. 需要训练一个与策略模型大小相当的价值模型（Critic Model），这带来了巨大的内存和计算负担
2. LLM 通常只有最后一个 token 会被奖励模型打分，训练在每个 token 上都准确价值函数难
3. GRPO 避免了像 PPO 那样使用额外的Critic Model近似，而是使用同一问题下多个采样输出的平均奖励作为基线

# GRPO的优点

- 无需额外的价值函数：
  - GRPO 使用组内平均奖励作为基线，避免了训练额外的价值函数，从而减少了内存和计算负担。
- 与奖励模型的比较性质对齐：
  - GRPO 使用组内相对奖励计算优势函数，这与奖励模型通常在同一问题的不同输出之间进行比较的性质相符。
- KL惩罚在损失函数中：
  - GRPO 直接将训练策略  $\pi_\theta$  和参考策略  $\pi_{ref}$  之间的 KL 散度添加到损失中，而不是像 PPO 那样在奖励中添加 KL 惩罚项，从而避免了复杂化  $A_t^i$  的计算。

## 5.DeepSeek-R1中尝试的失败的慢思考的案例

### Unsuccessful Attempts

- Process Reward Model (PRM)
- Monte Carlo Tree Search (MCTS)

## 2. DeepSeek-R1-Distill-Qwen-1.5B模型下载和转MNN

## 1. 下载DeepSeek-R1-Distill-Qwen-1.5B

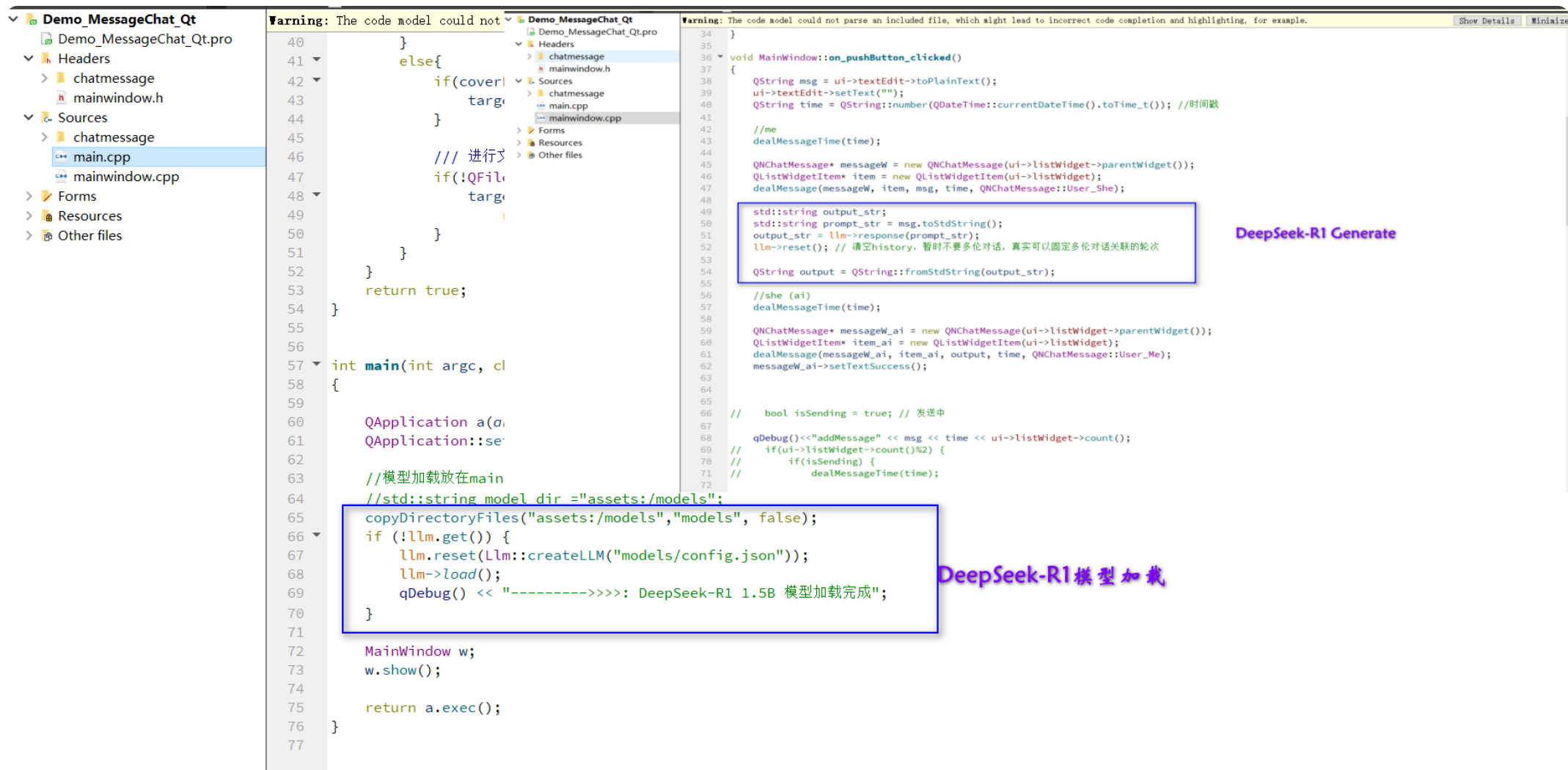
- GitHub: <https://github.com/deepseek-ai/DeepSeek-R1> ↗
- HF下载地址: <https://huggingface.co/deepseek-ai/DeepSeek-R1-Distill-Qwen-1.5B> ↗

## 2. DeepSeek-R1-Distill-Qwen-1.5B转MNN

- 按照我之前的教程编译安卓版MNN,并将DeepSeek-R1转为MNN: [https://dataxujing.github.io/Qwen2-0.5B-Android/mnn\\_qt\\_llm\\_slide.html](https://dataxujing.github.io/Qwen2-0.5B-Android/mnn_qt_llm_slide.html) ↗

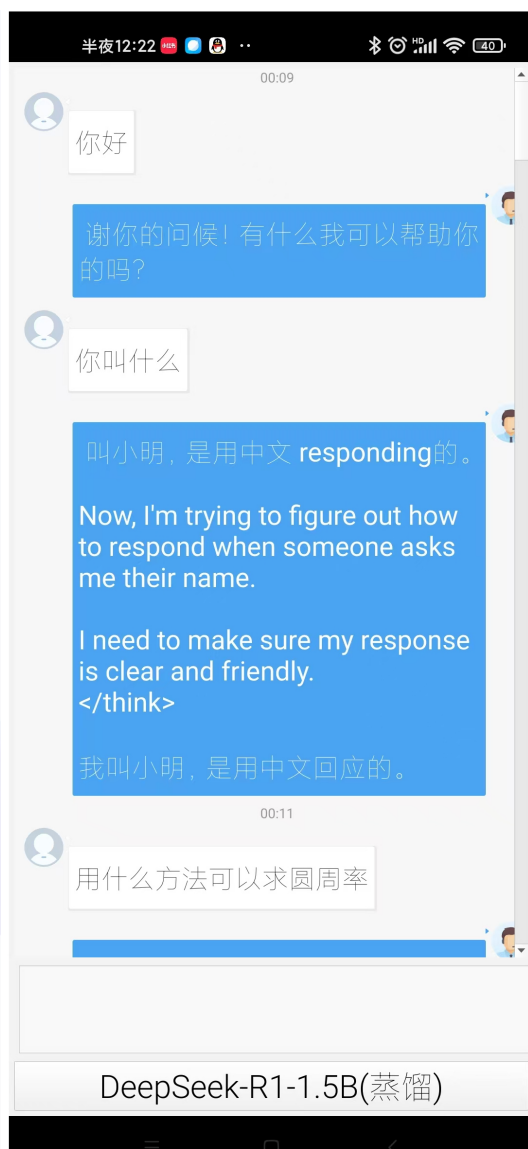
```
python llmexport.py \  
    --path /deepseek-r1 \  
    --export mnn \  
    -quant_bit 4 --quant_block 128
```

### 3. 基于QT和MNN的DeepSeek-R1安卓端部署





## 4. Demo



**The End**  
**Thanks For Your Attention!**