1. qwen32B 上pe，验证gametopic分类性能比较好（~96%）——pe都在user prompt上调优
   1. qwen8b的pe效果 （~94%）
2. 在外部输入的function desc上做pe调优，这块的效果是怎么样的
3. sft + function desc 调优，id效果（gameTopic在id验证集，在训练集）
4. sft + function desc 调优，ood效果（gameTopic在ood验证集，不在训练集）

# 问题

1. 意图分类场景的问题，新增类目（层次分类情况更复杂）往往会影响类目体系和现有类目的分类边界。在传统深度学习场景这里会涉及的现有样本标签修正的工作；而在LLM的PE范式下则会涉及到prompt的重新调整或重构
2. 新业务的模型模块往往会存在一个冷启动阶段，如新上线的领域智能体内部的意图分类，缺乏相应的样本（需要通过样本挖掘或合成数据手段缓解），会拖慢首期的项目迭代节奏。

# 动机

能否借鉴function call的工具调用任务形式，让LLM学会关注通过function description中的文本描述，对新的分类任务有一个较高的性能指标，以实现：

1. 能够以较小的成本实现冷启动阶段。
2. 能够比较好的隔离开各个类目的迭代，即降低 新增类目 或 优化特定类目时，对现有其他类目分类效果的影响。

注：此处的function description会包含大量的分类示例或边界描述，相当于把一部分prompt的优化迁移到function部分，类目发生变化时我们只需要优化相关的function description

# 难点

1. xxx

# 解决方案

1. xxx

# 效果

1. xxx