

# First Meeting Record

---

06/06/2023

## 1st 大模型方向

大模型的训练比较少，目前能做GPT3.5的prompt tuning（无法做

级联 不同的模型，链接及改进model performance

- 图像处理的模型
- 如果3.5可以实现，不需要用4（只能提升表现）
- visual GPT 可能不开放
- BLIP2：传言和GPT4的实现一致，图像理解能力可能是内部实现的training

## 2nd 小模型方向，没做过

也有可能做小模型 - 具体的task：小模型可以实现部分大模型的效果

虽然不能general，但是还是可以发paper的

文本转语音的技术

最终的benchmark 还是大公司主导

被秒掉怎么办？虽然被秒了但还是能立论写论文

重点在于开源：各个领域的expert，可以finetuning

## Time is limited

---

一个月是非常少的，具体一个大任务挑战很大

但是根据context 的语音场景的交流方式会更聚焦

有几个备选的方向，或自己提出框架下兼容的方向

第一次做海宁的暑研，没试过，看一个月内的成果

LLM + 具体的case

hyy: 行人, 非机动车的意图预测 -> 提醒人关注道路以外的

一个月很难出论文, 只能出论文中的一个模块

HCI: **Interactive NLP**

和环境交互的, 投的会议偏HCI

会议:

8月15, 一年四次, 12月有

9月can

期刊: 月刊or半年刊

7月初的C会

8月 triple AI, 专注于novel 的NLP

## 车内模态:

---

文字转语音 - 状态转换

文本挖掘 (数据挖掘, 文本处理) - 如何summarize

text2speech

比较传统的NLP 主流方向

- 机械性 or 情感性的
- 

## Corner case

---

自动驾驶的corner case: Stanford corner case 上个月发了

- ☐ 看论文, 查一下可以补充的部分
- ☐ CV的模型

## 轨迹规划

---

用LLM 解释routing, 分析如何做到传统算法的可解释性

自车的可解释性 ->

只用LLM

## resource

---

llm的论文不多，基本上是在做的工作（就两个学长学姐）

自动驾驶：感知部分的建模，互分析

对话代理模块？

需要看paper

- 租不到100，租3090，但需要抢
  - 云服务器做？Google
  - 部署也就20g，3090 跑~2天
  -
- 
- 定题后就知道需要其他的SOTA模型
  - 快速定题再分析模态的相关高被引paper
  - 光纯粹的模型部署是不一样的
  - HYY：视觉处理相关model
  - WJY：数字人，text2speech
    - 最重要：各种transformer model 的部署
  - 就一个想法：部署现有的模型，把pipeline的内容跑通（研究生水准了，还很顺）