

# 2022年终总结

Rui Hu

2023.1.7

# 研究方向

- 所在小组研究方向：**模型测试调试**，这个方向的目标是使深度学习模型能够和传统软件一样进行“测试”和“调试”

- 测试：发现模型问题
- 调试：修复模型问题



- 个人目前研究课题：**视觉模型去偏见**——解决因数据偏差导致的模型学习虚假相关性的问题，主要关注的是无监督去偏见

# 工作总结

- 研究论文

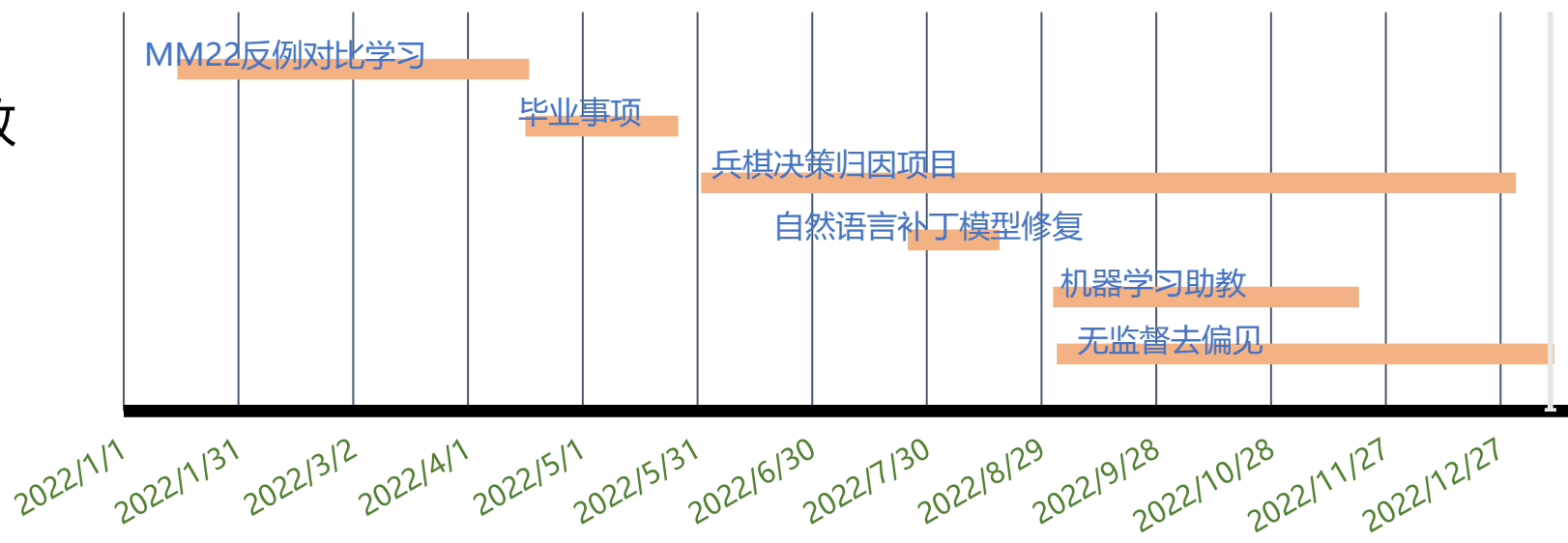
- 面向多偏见的模型去偏，**准备中**
- 基于反例对比学习的模型虚假相关性缓解，**已发表**MM22，辅助金强

- 技术项目

- 兵棋决策归因项目，第一版**已上线**

- 其他

- 22秋机器学习助教



# 入学至今工作量总结

- 积分：2
  - 素问前期项目 +1
  - 辅助MM22论文 +0.5
  - 兵棋决策归因项目 +0.5

# 2023工作计划

- 面向多bias的模型去偏见
- 视觉模型的特征学习优先级研究
- 基于自然语言补丁(多模态)的模型修复

