

高杨帆



教育背景

同济大学软件学院 工学学士

2019.09 - 2023.07

- 专业：软件工程
- GPA (前 6 学期): 4.51/5.00(3.61/4.00), 排名 78/212

项目经历

食神：基于 InternLM2 的菜谱百科全书

2024.02-至今

- 项目简介：
本项目是我参加书生·浦语大模型实战营的大项目，基于上海人工智能实验室开源的 InternLM2-Chat-7B 大语言模型，采用 XiaChuFang Recipe Corpus 提供的 1,520,327 种中国食谱进行微调，并建立外挂知识库，以利用 RAG 增强模型检索菜谱的能力。我们还在模型 pipeline 中增加了基于 SD 的生成菜肴示意图片的功能。我在其中负责 RAG 模块代码编写和调试。本项目入围书生·浦源大模型挑战赛 50 强名单，并作为优秀项目收录进官方社区。
- 项目代码：
<https://github.com/SmartFlowAI/TheGodOfCookery>

IMelodist：基于 InternLM2 的音乐生成大模型

2024.02-至今

- 项目简介：
本项目基于上海人工智能实验室开源的 InternLM2-Chat-7B 大语言模型，我们选择了基于 ASCII 码的乐谱记录方式，因此无需对模型进行多模态扩展就可以完成音乐生成的任务。目前，我们完成了 InternLM2-7B-chat 在 43 万条乐谱数据的增量微调和 SFT，并入围书生·浦源大模型挑战赛决赛答辩环节。我在其中主要负责微调数据集制作。
- 项目代码：
<https://github.com/GuoYiFantastic/IMelodist>

本科毕业设计：基于多种技术栈的聊天机器人

2022.07-2023.05

- 项目简介：
本项目为我的本科毕业设计项目。该项目利用 Rasa 聊天机器人搭建框架开发智能家居控制和车载语音助手两种场景下的聊天机器人，并使用 Transformers 微调了一个党史领域的问答机器人。
- Rasa 训练语料生成框架代码：
<https://github.com/SchweitzerGA0/rasa-train-generator>
- 专业领域知识问答机器人代码
<https://github.com/SchweitzerGA0/awesome-chinese-chatbot>

nlp-beginner

2022.12-2023.03

- 项目简介:

本项目是复旦大学邱锡鹏教授发布在 GitHub 上的针对 NLP 新手的系列任务，是我早期学习传统 NLP 知识的实践项目，实现了文本分类、NLI、NER、NLG 等传统 NLP 任务。

- 项目代码:

<https://github.com/SchweitzerGAO/nlp-beginner-impl>

- 实现日志:

<https://schweitzergao.github.io/2023/01/25/nlp-lab/>

实习经历

同济大学软件学院可信网络与容错课题组 2023.06 - 2023.08; 2024.01 - 2024.03

- 工作描述:

实习期间，我负责开发一套同声传译 APP，这个 APP 支持用户语音输入，在多种语言之间相互翻译并通过语音输出。此软件的一个工程难点是将语音识别模型和语音合成模型部署在端侧运行，我采用了 ONNX 作为这一问题的解决方案。

同济大学破壁工作室程序组

2022.01-2022.07

- 工作描述:

实习期间，我负责校园生活小程序“济星云”的 Java 后端开发工作。

技能

编程技能

Python, Pytorch, Transformers, Java, C++

外语技能

- 英语 (CET-6 601/750)
- 日语 (同济大学 2020 级日语辅修专业毕业)

所获荣誉及奖励

- 书生·浦语大模型实战营优秀学员
- 2022 年 MathorCup 全国大学生数学建模挑战赛一等奖
- 同济大学校三等奖学金 (2019-2020)
- 华特奖学金 (2020-2021)
- 同济大学优秀学生 (2020-2021)