## Imperial College London

# 姜禹尧



### ★ 教育经历

英国帝国理工学院

应用计算科学工程

2022.10 - 2023.11

- 硕士学位、专业前15%
- 核心课程:深度学习,机器学习,并行编程,运筹与优化,C++,Python

南京信息工程大学,英国雷丁大学

数学与应用数学

2018.09 - 2022.06

- 学士双学位, 专业前10%
- 核心课程: 概率论, 概率与统计, 数学分析, 高等代数, 抽象代数, 数值分析, Matlab

#### ■ 实习经历

• 中电鸿信信息科技有限公司

算法实习生(自然语言处理方向)

2023.06 - 2023.09

- 主要工作:研发自然语言处理(NLP)算法,识别并标记文本数据中的实体类型,提取不同实体间的关系, 从业务角度处理数据、调整标签、提取特征、训练优化模型、评估模型等
- 技术点:项目一采用两种方式(1.微调预训练模型Bert 2.微调魔搭平台的地址结构化要素解析模型)识别并标记地址信息中的实体类型,项目二采用通用信息抽取模型抽取任免文件中的实体和实体间的关系
- 产出: 地址实体识别模型和任免信息提取模型分别被江苏电信公司的业务受理系统和人力资源管理系统采用,地址实体标记的准确率和提取的任免职文本字段识别的准确率分别达到96%和99%

• 信也科技

#### 算法实习生(大模型方向)

2023.12 - 2024.02

- 主要工作: 1.根据公司的业务数据,在MoE, Llama2, Gemma等开源预训练模型上采用多个框架(TRL, LLaMA-Factory, Firefly, FastChat)进行**SFT微调**和**DPO训练**,实现多轮对话的推理并部署客服大模型 2.使用大模型做端到端的**对话管理**,应用**RAG**在预设的话术库中实现智能客服的话术选择和话术生成
- 技术点: Fine-Tuning, Quantization, RerankerModel, RLHF, 部署大模型推理的api及web界面
- 产出: 1.进行大量的微调实验以及调整learning rate, lora rank等超参数,在生产环境中上线运营反馈最好的客服多轮对话模型 2.基于RAG思路,独立编写对话管理的代码,用BCEmbedding包对用户输入和预设话术库中的问答对进行相似度对比,召回前top-k个问答对,调用大模型的api基于此选择话术或生成话术

#### 季 项目经历

● "概率和统计"主题的研究项目

#### 英国牛津大学

2021.06 - 2021.11

- 主要工作:在选定Student Performance数据集上应用随机森林和交叉验证从32个特征中选出权重最大的4个关键影响因子,降维的同时保证了多元线性回归算法对学生成绩预测的准确率
- 产出: 结项成绩84.18/100, 产出两篇论文, "Random Forest Regression for Predicting Student Performance" 和 "Statistical Analysis of Several Factors in Predicting Student Performance"

#### ☆ 公开成果

- 卷积神经网络算法研究 于2018年4月在《文理导航》期刊上发表,讲述CNN基本框架、实现原理及应用场景
- Statistical Analysis of Several Factors in Predicting Student Performance 作为CSAMCS 2021会议论文于2021年11月在SPIE期刊上发表,被EI,Scopus检索

#### ♥ 专业技能

- 掌握C++, Python编程语言, 熟悉PyTorch深度学习框架, TRL、LLaMA-Factory等大模型训练推理服务框架
- 熟悉传统的自然语言处理任务,在大模型微调技术和客服模型、多轮对话管理等下游任务的应用有实战经验

## ♥ 获奖情况

- 2020年全国大学生英语竞赛三等奖, IELTS 6.5, PTE 71, 大学生英语四六级通过
- 2022年校级优秀毕业生,多次校级三好学生,校级一等奖学金