

何浩坤

15992625392 | haokun.he@uzh.ch
广东省珠海市香洲区尚东领御13栋2单元1601

教育经历

苏黎世大学	2023年09月 - 2025年07月
人工智能 硕士	苏黎世
• 相关课程:人工智能前沿课题研究, 机器翻译, 自然语言处理中的机器学习算法, 数据库等	
澳门大学	2019年08月 - 2023年06月
计算机科学 本科	澳门
• 获得四次科技学院 "the Dean Honor List"	
• 相关课程:数据结构与面向对象编程, 机器学习, 模式识别, 多媒体编程, Android开发等	

实习经历

深译信息科技(珠海)有限公司	2024年06月 - 至今
NLP技术实习生 研发部	珠海
项目名称:多模态消防大模型构建与优化	
1. 项目目标: 为某市消防队构建多模态消防大模型, 主要功能为实时辅助消防救援, 消防知识科普等, 实现实时语音、图像、视频等数据的理解与传输。第一阶段专注于实现语音到文字的转录功能, 并针对消防救灾行动提供辅助决策支持, 同时辅助生成消防相关公文。	
2. 数据集预处理及向量知识库构建	
消防向量知识库建立: 基于预处理后的文本数据, 构建了一个高效、可检索的消防向量知识库, 且为消防场景的语义理解大模型微调提供了数据资源。	
结构化数据集: 针对客户提供的危化品相关数据集, 使用JSON来结构化危化品数据, 确保了数据结构的一致性和可用性。	
非结构化数据集: 对于纯文本或PDF格式等非结构化数据进行了整理, 通过大模型生成并构建了适用于消防大模型微调训练的问答数据集。	
3. 模型集成与调试	
应用实现: 采用LangChain/LangGraph框架, 将消防向量知识库与多种语义理解大模型进行了集成。	
模型调试: 通过精细化的调试, 确保了模型在理解复杂消防场景和语义上的准确性和效率。	
4. Rerank模型引入与优化	
重排机制:为了解决大模型在处理长提示语句时可能出现的注意力分散和信息处理不足的问题, 引入了Rerank模型。	
检索结果优化: 通过Rerank模型对向量知识库中的检索结果进行重排, 将精排后最相关的知识库文本放在prompt的最前面, 提升检索结果的准确性和相关性, 进一步增强了模型的实际应用价值。	

珠海格力金融投资管理有限公司	2023年04月 - 2023年07月
投资助理实习生-AI产业方向 金融投资部	珠海
1. 具体职责: 主要负责团队在人工智能行业的投资研究, 特别是分析ChatGPT浪潮影响下初创或行业领先NLP企业的产品及发展情况。	
2. 报告撰写: 完成了一份关于某小语种语音翻译及转录为核心的NLP初创企业的投资分析报告。通过与NLP行业专家和企业客户进行尽职调查, 并对比同行业竞争公司的数据, 详细分析了该企业在数据、算法及算力三大因素下的竞争力与投资价值。	
3. 深度交流: 独立与该初创NLP企业的CEO进行深入交流, 重点分析了企业在数据生成领域的可持续发展能力及未来产品的市场接受度。	
4. 投前估值: 参与商定企业的投前估值, 确保投资决策的科学性和准确性。	

华为技术有限公司	2022年06月 - 2022年09月
算法工程师实习生 终端BG HarmonyOS部	深圳
1. 任务需求: 参与完成“大规模细粒度菜肴识别”任务, 该任务要求模型识别图像中某种食物或多种食物的类型或名称。	
2. 基线网络构建: 为任务构建了三个基线网络, 分别基于CNN, Swin-Transformer和VOLO-Transformer, 并比较不同基线网络的识别效果。主要专注于构建VOLO-Transformer网络, 并根据模型预测结果进行数据集的调整和模型优化。	
3. 标准数据集建立: 为识别任务建立了一个菜肴数据集, 完成数据收集、数据清洗和数据标注等任务, 最终整体数据库包含超过	

- 760,000张图像和1,309个菜肴类别。
4. Demo程序开发: 构建了基于Android环境的Demo程序，将模型部署在端侧进行测试
5. 项目成果:最终版本模型Top5准确率到达97%，Top3准确率达到93%，产品已于2023年上线。

智行未来(珠海横琴)科技有限公司

2021年06月 - 2021年07月

数据分析实习生

珠海

1. 从各大旅游网站例如携程，马蜂窝爬取相关旅游数据，并用python，SPSS等相关统计工具进行数据分析;
2. 建立公司的云服务器及网站，分析了解不同类型云服务器的性能及适用情况，根据公司情况选择并建立相关云服务器，数据库及网站。

项目经历

CLEF 2024: Palöri at CheckThat! 2024 Shared Task 6: GloTa - Combining GloVe Embeddings with RoBERTa for Adversarial Attack

2024年03月 - 2024年06月

苏黎世大学

项目链接:<https://github.com/Muyuu77/InCredibIAE-2024-GloTa>

1. 任务目标: 此Shared Task的目标是设计一个攻击算法，针对多个特定领域的二分类任务，改变输入序列中的某个词汇形态(例如 "Water causes death" -> "Water cuases death"), 使预训练的二分类器对同样的输入序列输出不同的分类。 论文已被CLEF 2024接收。
2. 算法改进: 改进BERTAttacker算法，首先遍历输入序列，利用Masked Language Model (MLM) 计算每个词Masked和UnMasked对MLM模型输出的词概率差异，作为该词的重要性评分(Important Score)。
3. 替换词集合优化: 改进目标词的潜在替换词集合，使用RoBERTa模型作为基线模型，并结合Contextual Embeddings的信息，通过计算不同词之间的相似度，生成新的潜在替换词集合。
4. 结合GloVe Embeddings:在输入序列进入RoBERTa模型之前，先使用GloVe Embeddings来表示每个词，并设置阈值，在阈值内做词替换。
5. 性能提升: 将原先Attacker算法的BODEGA评分从0.42提升至0.54，特别是提高了攻击成功率。在测试集中，攻击成功率从290/416提升到400/416，最终在自动化检测和人为检测中均排名第四。

基于Wikipedia电影知识图谱数据集的Chatbot实现

2023年09月 - 2023年12月

苏黎世大学

项目链接：https://github.com/Muyuu77/ATAI_chatbot/tree/main

1. 此项目设计了一个可以理解自然语言及SPARQL语言的AI chatbot，主要针对基于wikipedia中的电影知识图谱数据集信息，对封闭/开放问题，电影推荐和多模态信息查询三种主要问题进行回复生成。
2. 对输入的问题做实体识别，并集成一个问题分类router，将输入问题分为以上三类
3. 对于封闭/开放问题，结合使用众包数据集中的数据及rule-based的方法，通过直接提取或计算实体embeddings相似度的方法来生成回复。
4. 针对电影推荐问题的需求，先对输入的电影样例信息进行知识图谱数据库查询，通过KMeans算法，提取出输入电影的共同特征，基于协同过滤的逻辑，返回给用户具有此特征的电影。
5. 同时基于内容的推荐算法，利用现有的电影embedding数据集，对于输入的演员、风格之类的信息，找到相应的电影embedding，通过embedding相似度来生成电影推荐。
6. 结合额外的MovieNet数据集，返回用户查询的演员、电影相关图片信息。

基于单个机器翻译模型的质量检测实现

2023年11月 - 2024年04月

苏黎世联邦理工学院

1. 实现NLP任务中模型自我质量估计(Quality Estimation)模块，重点是将有监督的QE直接集成到机器翻译模型中，以提高模型训练过程的可靠性和准确性。
2. 证明及利用监督QE来模拟人类评估分数，提高机器翻译模型最终输出的性能。
3. 分别开发集成了基于Token的方法和基于Regression分析的两种创新方法，用于机器翻译模型QE分数生成。
4. 实施包括基于蒸馏和self-QE方法在内的模型训练策略，进行全面的对比实验探索。

荣誉奖项

全国大学生统计建模大赛优秀奖-新冠疫苗对控制疫情的积极作用与后续不良反应现状分析

2021/06