

孙亚军

男 | 年龄：27岁

2年工作经验 | 求职意向：大模型算法 | 期望薪资：17-22K | 期望城市：武汉

个人优势

具有扎实的深度学习基础,在硕士三年中持续关注人工智能技术的发展,已熟练掌握现有大语言模型应用,遇到难以解决的问题时,积极向相关领域的研究者请教,并与广泛的研究人员和研究社区保存联系。

工作经历

武汉卓雅科技 数据分析师 2021.03-2021.07

负责对用户数据的收集与整理,选择合适的方法对数据进行可视化展示说明,从定性、定量和预测的方向,分析近期用户销售额的波动并给出可能的原因,

北京墨奇科技有限公司 算法工程师 2023.09-2024.05

- 针对公安部的多模态检索需求,我独立收集并制作了相应的数据集,随后训练了一系列多模态检索大型模型,以实现警方能快速准确地检索相关的图片和视频。
- 针对一些难以获取足够数据的特定检索场景,我采用了零样本学习的方法,成功实现了在数据不足的情况下仍能取得较好的检索效果。
- 我还将多个训练流程进行了工程化处理,使得无论面对何种新的检索需求,都能够迅速训练出一个符合要求的检索模型。并且,我将训练过程简化至用户界面形式,有效降低了研发成本,提升了版本迭代速度,
- 最后,我构建了多模态知识库,解决了现有技术仅支持单一模态知识库的问题,做出了让客户眼前一亮的AI产品;

某科技公司 AI大模型算法 2024.07-至今

2024.07-2025.01 多智能体项目

基于开源（QWen等）、闭源（盘古）大模型，构建垂直领域智能体。采用langgraph开源框架构建数据流、工具和反思，提高了可靠程度，部署在a100和华为310。

2024.10-2025.03 利用强化学习算法定制训练大模型

由于大模型存在幻觉，现有的数据不足以支撑较好的微调结果。采用强化学习算法训练出能够生成可靠的文本方案内容，基于强化学习生成的数据构建训练数据集，从而解决大模型幻觉。

2025.02-2025.06 基于langchain实现垂直领域问答需求（RAG）

基于稀疏和稠密向量方法，设计可变权重的检索方式，从而取得较好问答效果。

项目经历

新型信息数据检索系统 算法工程师 2022.02-2023.05

内容:

现有的检索系统如搜索引擎百度、谷歌,以及购物app淘宝、京东等均使用的是单模态建设方法,比如文字只能检索文字,图片只能检索图片,尽管淘宝可以实现文字检索图片,但背后是大量人工打标签的形式,原理仍然是单模态建设,随着多模态数据的增加,迫切需要跨模态建设技术以提高人们的信息检索能力,业绩:

主要完成算法的编程实现,搭建检索系统。

获得了案例分析国家一等奖,

项目内容:

该项目基于深度学习方法,在pytorch框架下进行多模态检索系统的研究。主要研究提高哈希学习效果的方法,以精确且高效的实现不同模态之间相互检索,并将其运用在日常生活领域,搭建新一代检索系统。

本人职责:

使用pytorch实现创新的算法框架,并收集收集进行模型的训练和测试负责模型的优化,降低模型复杂度,实现高效检索。研究新颖的可指定哈希学习的方法,并进行大量实验进行验证,

公安部多模态检索系统

大模型算法

2023.09-2023.12

·许多外国人抵达后未完成登记手续,目前的监测方法主要依赖于警方利用摄像头来辨识外国人并核实其登记情况。然而,这一方法的效率较低,导致大量未登记的外国人未能被识别,·针对这一问题,我收集并整理了相关训练数据,基于现有大模型构建了一个增强大模型。独立进行了模型的训练,成功地开发出了一种快速准确地识别外国人的产品,以帮助提高边境管理的效率和准确性。

静态场景下遛狗未牵绳行为的识别

2023.11-2024.05

·针对遛狗不牵绳的问题,这今为止尚未出现较为有效的方法来及时发现并提前预警。我的工作重点是训练一种能够识别未牵绳遛狗的大模型。我独立完成了前期数据的收集和处理、大模型的设计与训练,以及对模型效果的测试评估;

·通过训练,我成功地开发出了一种可靠的大模型,能够准确识别未牵绳遛狗这一违规行为,并及时发出预警。这一技术的应用大大降低了治安维护的成本,并提高了公共安全水平。

构建多模态知识库

2023.12-2024.05

国内公司推出的知识库通常只支持单一模态(文本),然而随着客户对知识库的使用增加,对于图像知识库的需求也日益增长、鉴于此,我进行了多模态大模型调研,发现现有的多模态大型模型能够更好地满足客户的需求。在此基础上,我独立进行了数据集的收集和制作,并对大模型进行了再训练。随后,我基于这一训练模型构建了特定场景的多模态知识库,从而研发出了一款市场领先的人工智能产品。这一产品的推出,满足了客户对于多模态知识库的迫切需求。

教育经历

湖北经济学院

本科

数据科学与大数据技术

2017-2021

广西科技大学

硕士

应用统计

2021-2024

硕士科研成果

已发表3篇中科院SCI论文

荣誉奖项

学科竞赛情况

第六届全国应用统计专业学位研究生案例分析大赛，全国一等奖；

② 第五届全国应用统计专业学位研究生案例分析大赛，全国二等奖；

③ “中国光谷·华为杯”第十九届中国研究生数据建模竞赛，全国三等奖；

④ 第十五届“中国电机工程学会杯”全国大学生电工数学建模竞赛，全国三等奖；

⑤ 第十八届中国研究生电子设计竞赛技术赛道，华南赛区三等奖；

⑥ 第十八届中国研究生电子设计竞赛商业赛道，华南赛区三等奖；

⑦ 其他奖项若干