

- ② 34岁
- ② 上海

Python

- = 10年经验
- 15236272162
- prwjb@sina.com

♂ 技能特长

Java

React

JavaScript

Shell

王家宾

◎ 教育背景

2010.08-2014.05

河南科技大学

本科

农学



工作经验

深圳百乐宝科技公司

2014.06-2016.04

FAE技术支持

- 负责对交付产品和服务的技术支持,一般性的技术问题咨询和解答。
- ■配合市场部门进行产品的宣传与推广及与客户沟通技术需求与方案。
- ■负责产品和服务在客户端的交付及应用培训。
- ■负责客户端的技术需求的收集和归纳,配合市场部做好项目立项技术准备 工作。

2016.06-2020.08

上海中科新牛命科技公司

信息工程师

■ SaaS云平台服务后端开发

负责设计和开发SaaS云平台的后端架构,确保系统的高可用性、可扩展性和安全性。包括开发RESTful API、实现微服务架构、数据库设计和优化,以及处理大规模并发请求。同时,参与到自动化部署和持续集成/持续交付(CI/CD)流程中,以确保新功能的快速迭代和高效交付。

■ windows桌面软件端开发

负责开发和维护Windows桌面应用程序,涵盖从需求分析、用户界面设计到功能实现的各个方面。开发具有丰富用户体验的应用,同时注重系统性能优化和用户交互设计。此外,负责应用的版本更新、错误修复以及与其他系统或平台的集成工作。

■ 数据分析流程开发

负责设计和实施数据分析流程,涵盖数据的获取、清洗、处理、分析及可视 化的各个环节。确保整个数据分析流程的可扩展性和可靠性,并不断优化流 程的各个环节,以提高分析的效率和准确性。编写技术文档,记录流程的设 计思路和实现细节,并为团队提供技术支持和培训,帮助他们更好地利用数 据分析工具和技术。

2020.09-2022.01

华大基因

中级信息工程师

■ 基于机器学习算法开发异常值检测可视化系统

负责设计和开发基于机器学习算法的异常值检测系统,旨在识别和处理数据中的异常点。包括选择和优化合适的机器学习算法(如孤立森林、支持向量机等),并将其应用于实际数据中,以准确地检测出潜在的异常行为。另外参与开发可视化界面,使用户能够直观地查看和分析检测结果。该系统不仅需要具有高效的检测能力,还要易于使用,并能够处理大规模数据集,从而帮助用户做出更准确的数据驱动决策。

■ 配合IT部对IDC集群的数据、软件等资源管理

与IT部门密切合作,管理IDC(Internet Data Center)集群中的数据和软件资源。包括服务器配置管理、虚拟化平台的维护、数据备份和恢复策略的实施,以及网络安全的监控和管理。另外参与到资源分配、性能监控、故障排查等任务中,以确保IDC集群的高效运行和安全性。

高级软件开发

• 大模型应用开发 (NL2SQL、专利RAG、垂直领域知识图谱等)

负责开发基于大语言模型的应用程序,如自然语言转SQL(NL2SQL)、专利检索及生成(RAG,Retrieval-Augmented Generation)和垂直领域知识图谱。参与从概念验证(POC)以及最终交付,涵盖从数据收集、建模到系统集成的整个过程。另外负责优化模型的性能,确保其在处理大规模数据时的响应速度和准确性。

■ AI 制药 SaaS 平台前后端全栈开发

主导AI制药SaaS平台的全栈开发,涵盖从前端用户界面到后端数据处理的各个层面。使用现代前端框架(React)开发直观的用户界面,同时在后端使用Python等技术处理复杂的AI制药算法。负责设计系统架构、管理数据库、实现API,并确保平台的安全性和高可用性。为制药行业提供强大的工具,帮助科研人员更高效地进行药物研发和数据分析。

■ 开源大模型在HPC上进行分布式训练SFT微调 本地和AWS 云计算平台推理 部署

负责在高性能计算(HPC)环境中进行开源大模型的分布式训练和SFT(Supervised Fine-Tuning)微调,以提高模型在特定任务上的表现。使用诸如PyTorch Distributed或Llama-factory等框架来处理大规模模型的分布式训练,并优化训练过程中的资源利用率。此外,负责将微调后的模型部署在本地和AWS云平台上进行推理,确保其能够在不同的计算环境中高效运行,同时满足业务需求的性能指标。

■ IDC机房HPC搭建及HPC、Gitlab运维

在IDC机房中负责高性能计算(HPC)集群的搭建和日常运维。包括硬件资源的规划和部署,如计算节点、存储系统和高速网络的配置,以及HPC软件环境的安装和优化。管理HPC集群的作业调度系统,确保资源的高效利用和任务的按时完成。此外,负责HPC环境的监控、故障排除和性能调优,确保集群的高可用性和可靠性。

主导GitLab平台的从零搭建工作,包括服务器的初始配置、GitLab的安装与配置,以及相关的用户管理、权限设置和项目创建等工作。负责制定并实施数据备份策略,以确保项目代码和相关数据的安全性与可靠性。除了基础架构的搭建,并设置了CI/CD(持续集成/持续交付)流水线,自动化构建、测试和部署流程,显著提高了开发团队的工作效率。负责监控和优化GitLab实例的性能,定期更新系统以保持其安全性和稳定性,并提供技术支持和培训,帮助团队更好地使用GitLab进行项目管理和协作。



作品展示

个人博客

https://shubihu-github-io.pages.dev

个人网站

https://letsthink.top

- 些证书

https://shubihu-github-io.pages.dev/certificate