

孟晓

经验：5+

求职意向：数据科学家、NLP、算法工程师、AI Engineer



santochaoya@yahoo.com.hk



+86 17628072710



https://www.linkedin.com/in/xiao-meng-851b89153



https://github.com/XimeCraft?tab=repositories

自我介绍

具有 5 年以上数据科学、数据挖掘、数据建模、数据分析和生成式人工智能(GenAI)，专注于 金融、通信和工程咨询 等领域。

在自然语言处理(NLP)及大型语言模型(LLM)积累了丰富的实践经验，熟练掌握 HuggingFace Transformers、LangChain、RAG 及 Azure OpenAI等技术。

同时专注于时间序列预测(Forecasting)与分类(Classification)模型，精通 LSTM、ARIMA 及多种机器学习与深度学习算法，并运用于多个行业场景，熟悉并实践完整端到端深度学习项目开发生命周期。

工作经历

EPAM System (Singapore) 数据科学家 2024.03 - 至今
EPAM System 亿磐系统科技（成都）有限公司 数据科学家 2021.09 - 2024.03

- 构建端到端债券时间序列预测模型，利用时间序列分析、深度学习和机器学习
- 开发基于新闻的 NLP 解决方案，包括 Transformer 模型和 GenAI 应用
- 运用假设检验、时间序列分析、生存分析和 A/B 测试，进行市场统计分析
- 搭建 WebSocket 与 RESTful API 实现实时数据处理
- 实现金融数据的端到端数据管道，包括网页爬取和数据库集成
- 开发自动化建模和分析工具，以优化数据清洗、特征工程和模型评估

亚信科技(成都)有限公司 算法工程师 2021.05 - 2021.09

- 构建与优化机器学习模型，用于用户体验评估、客户满意度优化与风险管理
- 基于电信基站时序数据，开发定制化机器学习模型与多步预测算法，提升预测精度与异常检测能力
- 多个机器学习建模全流程开发，包括数据处理、模型训练与优化部署，涉及时序预测、异常检测、增量学习与用户满意度分析

Mott MacDonald (New Zealand) 数据科学家 2019.05 - 2021.01

- 数据挖掘与处理，集成 API、爬虫、AWS S3、GIS 平台、ETL 与云函数（Azure Function, CI/CD）
- 构建时间序列模型与机器学习模型，实现预测与分析
- 异常点处理、显著事件提取分析，结合 Power BI进行数据可视化与面板开发

四川省丝绸科学研究院 会计 2012.05 - 2015.09

成都优力微波通信有限公司 硬件设计、测试 2010.09 - 2012.03

教育背景

坎特伯雷大学(新西兰)

2017.07 - 2018.12

计算机科学(本科)

西南财经大学

2013.03 - 2014.12

会计(硕士)

电子科技大学成都学院

2006.09 - 2010.07

电子信息工程(本科)

技能

● NLP & Generative AI:

实体提取，实体匹配，文本分类，情感分析，文本处理, Transformer, BERT 和 GPT 系列模型，Generative AI, Prompt Engineering, RAG, FAISS, LangChain, LoRA, Azure Cognitive Search, GenAI 评估优化, Vertex AI

● 模型预测 & 分析:

时间序列预测, 分类预测, 时序模型 (ARIMA, LSTM, Prophet, Seasonal Decomposition), 机器学习, 深度学习, 数据挖掘, 数据分析, 统计分析, 数据可视化, 推荐算法, 异常检测, A/B 测试, 特征工程

● 机器学习工程:

Airflow, MLflow, Data Drift, CI/CD Jenkins, OpenShift, Docker, Evidently AI, ETL Spark, Azure AutoML, Azure OpenAI, AWS