

毛天翔

男/1991.6

✉ mtianxiang@gmail.com

🌐 <https://github.com/POFK>

📞 13041256081

🏠 更多信息及最新项目请见个人主页 <https://txmao.vip/about/>

个人信息

我的标签：[博士后](#) [快速学习](#) [问题解决者](#) [系统设计](#)

我的能力

- 快速学习** 能在一周内学习并使用绝大多数编程语言，能快速理解不同行业背景和原理
- 解决问题** 源自科研工作中积累的思维模式和经验，能短时间看透问题本质，定位堵点并给出创造性的解决方案
- 个人管理** 注重个人时间管理和项目管理，利用GTD与OKRs管理日常工作和生活

学术背景

- 博士后** 目前为中国科学院国家天文台博士后、特别研究助理
- 学术成果** 有4篇论文发表在天文领域顶级期刊APJ和MNRAS上（[点击查看](#)），总引用数31，并另有1篇专利和3篇文章在投
- 科研经历** 中国科学院“稳定支持基础研究领域青年团队计划”入选项目（2000万经费支持）成员、参与1项国家自然科学基金（63万）、2022年受邀作为3名主讲嘉宾之一参加未来论坛闭门耕“AI+天文”主题线上研讨会、2019年受邀对英国杜伦大学和爱丁堡大学进行学术访问
- 研究方向** 主要研究领域为宇宙学、星系形成与演化理论以及AI在天文中的應用等，另外曾作为重要成员参与射电干涉阵列数据处理分析（数据量级TB以上）

管理经验

- 有5-8人左右科研团队的管理经验
- 参与多个国际合作项目

教育背景

博士	中国科学院大学 - 天体物理	2016-2021
硕士	中国科学院大学 - 天体物理	2014-2016
学士	兰州大学 - 物理学	2009-2013

工作经历

中国科学院国家天文台 博士后、特别研究助理

2021.7 - 至今

聚焦“宇宙三维初始密度的重构与再模拟”这一国际重要问题，带领多名博士生发展出了一整套基于神经网络的数据处理方法，比较好地给出了一系列针对星系红移畸变修正、星系群证认以及密度场重构的算法。目前我们的多项结果处于国际领先。

洞见科技 算法顾问

2021.11 - 2022.6

作为算法顾问加入领先的隐私计算技术服务商——洞见科技，参与算法研究、开发等。

期间，针对公司隐私计算平台，主导了联邦场景下横向和纵向深度神经网络算法的研究和升级，帮助公司通过PoC。同时设计和开发了面向用户的联邦学习平台安全验证方法，该方法已发表专利。

项目经历

个人项目：基于特定数据的对话机器人和问答系统

2023.5 - 至今

ChatGPT 向量数据库 text embedding langchain react docker prompt 优化

ChatGPT的出色能力有目共睹，但其依然存在诸多缺点。为了增强gpt生成回答的时效性、相关性并同时削弱样本偏差和事实错误带来的影响，通过为其集成特定数据库和工具来限制gpt的内容生成是一个有前途也容易实现的方向。

本项目借助langchain等开源项目，尝试实现一个基于特定文本或数据库的知识问答系统。开发此项目的目的是为我的个人知识数据库制作一个智能助手，提高索引效率。

个人项目：借助 AI 大模型将艺术品画作转换为3D模型

2023.4 - 2023.5

机器学习 cv大模型 open3d UE5 docker

历史中留存下来的大量艺术品画作或照片往往具有非凡意义，但由于其复杂场景和只有单视角信息难以通过NeRF等算法很好地恢复3D模型。本项目利用视觉大模型对单张二维图像进行三维重建，主要适用于符合光影几何物理规律的画作（如大量文艺复兴时期的作品）或大部分照片。

具体demo效果请见 <https://txmao.vip/about/#project> ,点击图像可观看动画。

联邦学习平台安全验证方法

2021.11 - 2022.6

联邦学习 区块链 有限状态机 消息队列 流量复制 神经网络

本项目基于公司的联邦学习计算平台，目的是为用户设计安全验证方案从而使平台获取用户的信任。其难点在于平台用户往往不具备专业的机器学习和密码学知识，当其对平台缺乏信任时，极端情况会将安全验证工具与平台一同视为黑箱，从而对安全验证工具也失去信任。

本人作为算法顾问，负责整个项目的方案设计、开发、协调沟通等。在对具体场景进行调研和分析后，创新性的提出了多级化的验证方案，并在与CTO以及不同部门的沟通中逐渐完善方案并推动该方案被接受。该方案涉及的技术栈包括有限状态机模型、流量复制和解析、区块链持久化等。

在项目推进的过程中，还负责了公司计算平台的横向、纵向神经网络的设计、开发和平台集成，并帮助公司通过项目POC。

气体检测设备的软件控制系统

2021.9 - 2021.11

Flask Python Redis Docker 串口 软件控制 嵌入式

本项目设计和开发了一套气体检测系统，其中软件部分包括可视化控制面板、设备自检、自动观测、数据存储、传输等功能。

作为团队负责人，负责整个项目的方案设计、路线规划和项目管理，同时参与了软件开发、测试及部署。期间为了推动项目进展，对硬件选型进行了大量调研，并学习了串口控制、docker、k8s、前端、后端等相关知识。

- 测试驱动与敏捷开发
- Python实现串口通信与API开发
- Redis内存数据库加速及持久化
- Docker部署

基于AI的新一代星系巡天数据处理方法

2019.10 - 至今

pytorch DNN GNN 机器学习 数值模拟

星系巡天数据在未来十年将迎来井喷式的提升，但数据分析方法却还未做出相应的升级。如何将观测到的星系数据映射到更本质的暗物质分布是其中最重要的堵点之一。本项目开发了多个基于AI的方法来处理这一问题。

负责科学问题发现、方法设计以及博士生团队的指导和管理。过程中通过从问题中提炼分解出三个最重要的独立子课题，使多名博士生可以同步推进项目。目前本项目已经在多个方面取得了国际领先的结果，正在准备论文。

用神经网络重构宇宙初始状态

2017.9 - 2019.6

机器学习 Python TensorFlow 分子动力学模拟 三维卷积神经网络

本项目利用三维卷积神经网络从天文观测数据中对宇宙极早期物质分布进行回归重构，结果发表在天文领域顶级期刊MNRAS上，并被国外媒体多次报道。

负责科学设计、项目规划、数据清洗、模型训练以及结果分析。

天文数据并行分析软件 (CosmAna)

2016.9 - 2018.1

大数据 并行 傅里叶变换 Python Cython

本项目专注于解决Python在处理天文大数据中面临的性能瓶颈，主要用于并行计算功率谱及相关系数，将计算时间从小时加速到秒量级。Github地址为 <https://github.com/POFK/CosmAna>。

负责项目规划、算法设计、开发及测试。

- 测试驱动开发 (TDD)
- 基于mpi4py的MPI并行计算
- 利用并行的三维傅里叶变换进行梯度计算
- 利用C优化封装核心函数，完成计算时间从小时到秒量级加速

恒星光谱分类

机器学习 Python TensorFlow 分类 光谱

本项目利用多层卷积神经网络对恒星光谱进行分类，以处理未来大量的恒星观测数据，挖掘更多潜在信息。该项目被纳入多个国际望远镜合作计划，Github地址为 <https://github.com/POFK/starspectraclassifier>。

参与科学设计和项目规划，负责原型开发、模型训练等。

技能

Python 编程 主要技能 10年经验

- 熟悉基本语法、标准库以及高性能计算库
- 精通面向对象编程和单元测试
- 精通利用C以及并行编程对Python进行加速
- 用Python做过科学计算、机器学习等项目开发

机器学习 算法 主要技能 7年经验

- 精通PyTorch, TensorFlow 等机器学习工具
- 熟悉CNN、GNN、VAE等多种神经网络算法
- 做过5个以上机器学习项目
- 精通数理统计、贝叶斯分析等

主要编程语言

- Python *****
- Docker *****
- C ****
- MPI ****
- Lisp ***
- javascript ***

Docker & k8s 运维 主要技能

- 熟悉Docker、k8s架构及工作、方式
- 搭建并管理过k8s集群
- 日常工作环境为自己定制的远程桌面镜像，运行于k8s集群中
- 在自搭建的k8s集群中部署过 kubeflow 机器学习平台

OKR & 敏捷开发 管理 主要技能

- 用GTD管理个人日常工作生活
- 有用OKR和敏捷开发范式管理5-8人团队经验

- Golang *****
- k8s ****
- Linux shell ***
- git ****
- Redis ***
- CSS ***