



潘杰

出生日期: 2001.06.28 届别: 2027 届 电话: 153-7644-9841
学位: 硕士 政治面貌: 团员 邮箱: 15376449841@163.com

基本信息

2024.09-2027.06

浙江科技大学

环境与资源(机器学习交叉)

导师: 郑兴教授、杭州辰禹环境科技有限公司 总经理 王禹浪高级工程师、陈镒锋讲师

目标岗位: 智慧水务工程师、AI 技术工程师、时序算法工程师

研究方向: 智慧水务以及智能算法辅助下催化剂的理性设计

校内荣誉: 连续 2 年获得学业奖学金, 优秀学生会干部

实习经历

2026.07-2026.08

安道麦辉丰

盐城

智慧水务实习生

业务场景: 化工园区污水处理厂生产废水必须严格达标 (一级 A 排放标准 GB 18918-2002), 传统人工巡检以及日级化验分析在水量水质波动大时响应滞后, 水厂需要及时信息反馈与工艺稳定提升。

业务职责: 参与园区污水处理厂**日常工艺巡检**、负责 SCADA **数据采集归档**、辅助进出水水质化验与异常排查。

业务内容: 每日巡检预处理单元、生化池、二沉池、深度处理单元, 记录工艺参数异常并上报。对接厂区 DCS/SCADA 系统完成连续多月的进水流量、COD、NH₃-N、TP、pH 等关键指标的批量导出与**结构化整理**。配合化验室完成**进出水水质日检**, 整理化验记录做趋势分析。参与 2 次**出水指标异常溯源**, 通过回溯历史数据定位到上游生产车间间歇性排水。

业务产出: 从数据产生端完整理解了工业数据质量常见问题如: 传感器漂移、通信中断、化验频次不足, 为**智慧水务**算法建模提供了**工程视角**的输入。

实习经历

2023.06-2023.08

网易数字产业基地

南京

python 实习生

业务场景: 运营团队每天需手工从多个电商平台导出销售、库存数据, 日均处理约 **8.5 万**条数据, 人工操作耗时每天约 1-1.5 小时, 经常出现**遗漏或延迟**等问题, 需要通过自动化数据替代手工汇总, 确保数据的完整性与实时性。

业务职责: 参与自动化日报系统的开发与定时部署, 包括数据清洗、报表生成、定时推送、异常告警。

业务内容: Pandas + Requests 定时从电商平台 API 获取销售、库存、退款数据, 清洗合并后入库。Openpyxl 自动生成格式化 Excel 日报。配置 Linux Crontab 定时任务 和 企业微信 Webhook 告警。编写 **27 个 Pytest** 单元测试。

业务产出: 日均清洗 8.5 万条数据, 每日 9:00 准时推送日报至运营群。为**智慧算法**建模提供代码**工程基础**。

项目经历

2026.04-2026.07

智慧水务 agent 开发

业务场景: 污水厂工艺异常诊断需**严格依据国标** (GB 18918 等)、**运维案例**等知识, 传统方式依靠个人经验积累, 新人培养周期长。需要将国标要求、历史案例与专家的经验沉淀, 形成可检索的知识库, 用于指导新人开展异常识别与工艺调控。

业务职责: 负责 Agent 系统架构设计与核心模块开发, 涵盖水务知识 RAG 检索管线、安全 Agent 图、MCP 协议接入与 Docker 部署。

业务内容: 基于 LangGraph 构建安全 Agent 图, 检索结果经引用门控校验后输出, 确保每条结论附带国标、导则、论文等来源依据。数据层以 **Qdrant** 向量数据库支撑 **BM25** 与稠密检索混合排序, 经 **reranker** 重排, 覆盖工艺参数、排放标准、运维案例等水务知识。通过 **MCP** 协议暴露 3 个只读工具, Docker Compose 一键部署, 测试覆盖 **agent、API、data、RAG** 等七个模块。

业务产出: 将分散的水务知识整合为可检索的 RAG 知识库, 加快工艺异常排查, 拓展了**智慧水务**在企业的应用。

论文成果

智慧水务：小时级进水水质预测模型(第一作者|2025.10-2026.07|浙江径山|6744 条小时级数据|已完稿)

题目： Leakage-Safe Rolling Evaluation for Hourly Influent Forecasting in a Full-Scale Wastewater Treatment Plant

业务场景： “十四五”智慧水务规划明确要求推动城镇污水处理厂向智慧化运营转型。在“双碳”目标下，精确加药与精细化调控成为降低能耗药耗的关键路径。然而，目前多数水厂仍依赖日级人工取样和化验室分析，异常发现滞后 **12-24 小时**，极易错过最佳工艺调控窗口。尤其是 $\text{NH}_3\text{-N}$ 出水超标，单点即可能触发环保处罚，给稳定达标运行带来较大压力。

业务价值： 模型准确度提升 **60%+**，95%部署级预测区间宽度较普通 **bootstrap 扩大 15-91%**，量化预测不确定性。模型长期运行 R^2 稳定在 **0.970-0.978**，降低运维成本。

技术栈： scikit-learn · PyTorch · Dlinear · PatchTST · Ridge · iTransformer · Block Bootstrap

论文成果

智慧水务：日尺度灰箱模型(第一作者|2026.03-2026.06|河北秦皇岛|1065 条日级数据|已完稿)

题目： A grey-box STL-RF framework for multi-horizon forecasting of influent total nitrogen in wastewater treatment plants

业务场景： TN 是污水处理厂最难稳定达标的指标之一，单次超标即可能触发环保处罚甚至停产风险。依据《城镇污水处理厂污染物排放标准》(GB 18918-2002) 一级 A 标准, TN 限值为 15 mg/L。本文进一步参照 BSM2 (IWA Benchmark Simulation Model No.2) 国际基准框架，对预测结果进行工艺响应评估与影响分析。

业务价值： H=3 天 $R^2=0.990$ / H=7 天 $R^2=0.919$ / H=14 天 $R^2=0.849$ ，**中长期步长优势更显著**。基于 BSM2 基准工艺环境的响应分析有助于识别关键调控窗口，为水务集团制定**工艺调控标准化流程**提供参考依据。

技术栈： STL (LOESS) · Random Forest · TCN · SHAP · BSM2 · XGBoost

论文成果

智慧水务：信号分解的日尺度预测模型(第二作者|2025.10-2026.04|浙江南浔|1461 条日级数据|已投稿)

题目： A hybrid multiscale decomposition and Informer-based residual learning framework for influent quality forecasting of wastewater treatment plants

业务场景： 单一模型难以兼顾长序列**预测的精度与稳定性**：传统 ARIMA 处理非线性弱、Transformer 在小时级波动上易过拟合；工艺工程师既需要**长期趋势**也关注**短期异常**。

业务价值： 在 3 d 预测步长上 R^2 达到 0.990，14 d 步长上仍保持 0.849，能够为**每周的调度规划**提供支持。残差诊断工具 (ACF 与 FFT) 进一步识别出约 2.0 d、2.8 d 和 3.8 d 的亚周振荡周期，揭示了传统建模方法容易忽略的固有周期模式。

技术栈： VMD · ARIMA · Informer (ProbSparse) · LSTM · ACF · FFT

论文成果

高精度能量预测与可解释公式探索(2026.03-2026.06|3424 个材料结构样本|已投稿)

题目： Machine learning-assisted rational design of high-entropy perovskite oxides for efficient oxygenevolution reaction

业务场景： 在高通量材料筛选中，传统 DFT 计算单个结构约需 24 小时，面对 3,424 个结构及数万个候选体系时，总计算时间超过 80,000 小时，完全不可行。为替代这一昂贵流程，本研究系统对比了四类**图神经网络架构** (AOR-GNN、CEGANN、ABSGCNN、ALIGNN) 与传统机器学习方法，在 3,403 个材料结构的畸变二分类任务上进行标准化评估，其中正类样本仅占 8.84%，构成显著的不均衡学习难题。

业务价值： GNN 推理模型将单结构预测时间从 CPU 小时级压缩至秒级，**大幅提升筛选效率**；同时，采用 SISSO 符号回归输出可解释性公式，并结合 VOR 算法与领域知识构造融合特征，经降维方法对比后，通过 RFE 构建最终特征集。ABSGCNN 在验证集与测试集上的 **AUC 均超过 0.92**，验证了该框架的有效性。

技术栈： CatBoost · LightGBM · ASE · pymatgen · catlearn · SISSO · GNN

证书技能

英语四级证书、普通话证书

自我评价： 7 个工业级项目覆盖**智慧水务 + 材料 ML + AI 工程**三个方向。开发 AquaOps Copilot 基于 **LangGraph + Qdrant** 的水务知识 **RAG** 与安全 Agent, 以 MCP 协议暴露只读工具, Docker Compose 一键部署。对数据泄漏审计 (四协议对比)、时序自相关校正 (ACF+FFT 残差诊断)、不确定性量化 (共形预测)、Agent/RAG 工程化 (LangGraph + 向量检索 + 引用门控) 有系统性实践。每个项目从数据清洗、模型训练到工程部署闭环，具备从**模型到产品**的端到端交付能力。