

某水务集团水质监测 AI 时序预测可行性验证

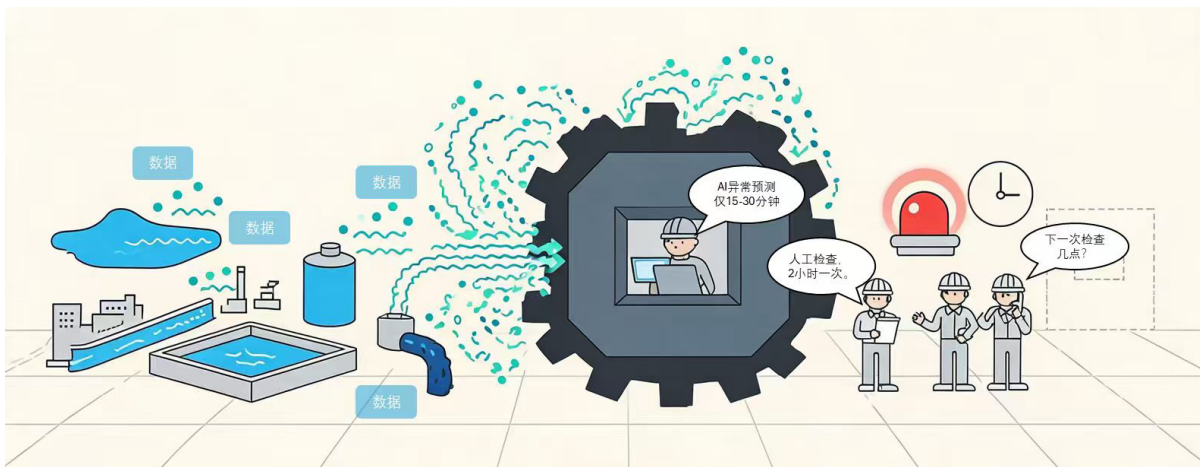
案例#: AI_TT_2026002

某国际水务集团的城市公共水体水质监测场景中，我们从大语言模型 Transformer 架构本质出发，利用企业自有数据，将其转化为 AI 可学习的标准化数据，并自建时序预测模型，将水质异常告警从平均 30-120 分钟（人工巡检发现）提前到仅需 15-30 分钟（AI 预测识别并预警），为水质安全与人工干预争取关键处置窗口。全程处理本地化，数据不出域。

场景综述

水务行业是典型的公共事业和重资产行业。地市级水务集团通常管理着数百个水质监测点（涵盖水源保护地自然水体、城市内公共人造水体，到自来水厂、二次供水、污水排放等水质检测区域），每天产生大量的人工巡检和已经数字化的物联网（IoT）水质数据（PH 值、水体温度、悬浮物混浊度、导电率、藻类、细菌等）。日均数据量常有数十万条。

水体水质预测是水务集团的日常核心运营需求：水质异常预警、工艺调度优化、人工巡检替代、合规报告自动化。然而，传统基于自回归积分移动平均模型（ARIMA）、指数平滑的时间序列统计模型，对多变量、非线性、突变型水质变化预测能力有限，大量 IoT 数据沉睡未用，异常发现依赖值班专家经验判断，响应延迟通常达 30-120 分钟，可能导致错过最佳处置窗口。



真实痛点

经 FDE 团队与水务集团运营团队沟通后，了解到两个亟待解决的痛点：

- 现有系统主流仍是传统统计学模型，数据处理量基于人工巡检的收集能力，大量 IoT 水质数据未充分利用，导致预测能力有限。当前大部分 IoT 实时采集的水质数据仅用于基础统计，大量信息线索未被挖掘，异常告警依赖值班专家人工判断，响应延迟平均达 30-120 分钟。
- 响应十五五“加快高水平科技自立自强，引领发展新质生产力”政策，探索充分挖掘 AI 替代重复性人工巡检。客户决策层明确诉求：利用 AI 技术升级水质运营，从“经验判断”升级为“AI



预测+人工复核”，让历史 IoT 数据产生真实业务价值，提升数字化投入的真实 ROI 水平。同时，在不增加人力成本的前提下，真实提升水质安全管理水平。

FDE 视角与洞察

客户最初提的核心诉求是“提升水质异常告警准确率，缩短发现到告警的时长”。但真正蹲到现场才会发现，告警准不准只是表象，能否从事后告警升级为事前干预才是目的。一个是预警模型表现再好，它只能“等水质变了再告警”，它只是把问题从“30-120 分钟后发现”变成“15 分钟后发现”——时间缩短了，但响应模式没变，本质仍是事后应对。

真正的价值是“提前预判水质会发生什么，在污染扩散前主动干预”。拿天气预报做个类比，这相当于不是简单地在下雨时发现下雨，并用最快的速度通知相关人员；而是能否预测明天下雨，然后建议安排事先人工降雨，避免明天的关键时间点有下雨的影响。

换言之，水质异常告警准确率只是一个过程技术型指标，真实的业务需求在于：让 AI 不仅能“及时看见”异常，还能“高效预测”异常，让水务集团从“等雨来”升级为“提前防雨”。

解决方案

我们围绕“可持续学习的数据预测能力”这个核心来构建 AI 预测的试点验证（POC）方案：

与传统的基于统计学的时序模型不同，大语言模型的 Transformer 架构天然拥有在序列预测上的核心优势，且这种能力已经在自然语言处理的长文本上获得了性能与工程化能力的事实验证。大语言模型本质上是精准预测一长串 Token 的预测模型，通俗地说，就是“快速准确地猜出下一个字”的算法。在我们这个场景里，我们要“猜”的是下一个水质向量（Token）——这正是 Transformer 时序预测的核心能力。

具体的试点验证方案，分为以下步骤：

从业务知识解构水质数据，找出影响水质的核心指标（此类指标一般符合幂率分布），并将水质数据点向量化，成为可训练的 Token。这个准备过程，类似在大语言模型中待训练的语料集。

1. 参考 GPT-2:124M 的 Transformer 架构，对向量化私有水质数据从新预训练。
2. 使用经业务团队标注的水质数据集，对步骤 2 生成的预训练模型进行分类微调。
3. 实测 124M 小规模参数的基于 Transformer 架构对水质数据的预测效果和泛化能力。

同时，对以下两个辅助方案进行落地的工程化探索：

1. IoT 水质预测传感设备的数据归拢与机器学习标准化自动处理。
2. 非接触式水质预测。利用普通摄像头拍摄的露天水体水面颜色，配合多模态大模型和历史人工干预处置事件记录，在不同水质预警场景中灵活适配误差容忍度，作为 IoT 传感器之外的冗余补充预测手段，提升整体预警覆盖率。

上述三个试点验证研究，跳出了仅使用大模型在文档处理方面的应用，选择直接从 Transformer 架构和神经网络预测算法层面入手。直接预训练出一个参数规模适当的专用大模型，通过自有数据



标注对预测后的水质数据做分类微调，据此预警业务上是否需要人工干预处置。IoT 与非接触式水质预测，则专注在新增水质数据的持续自我迭代和预测及时性。

最终的落地方式：

1. 一套本地工程化部署的 Transformer 预测服务：基于 PyTorch 从零实现多头自注意力、位置编码、训练循环，既保证预测精度，又避免数据出域。
2. 一个数据闭环机制：半自动使用新增水质数据做增量训练，模型权重随业务一起迭代，真正成为“会自己长大的 AI”。
3. IoT 硬件传感设备的试验性搭建与数据采集验证。

实施效果

本方案以“可持续学习的预测模型”为核心目标，从单点水质异常告警需求切入，设定明确改进指标：让水质异常告警从“事后发现+告警”升级为“事前预测+预警”，并在业务运行中持续自我迭代。

改进效果从三个维度衡量：

1. 水质数据时序预测 – 技术指标

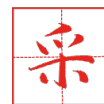
指标	改造前	改造后
水质数据预测准确率	/	~ 80%
水质数据预测单次耗时	/	< 3 分钟
异常告警误报次数	平均 20 – 30 次/月	平均 5 次/月

2. 水质运营 – 业务价值指标

指标	改造前	改造后
异常告警提前量	平均 30-120 分钟	15 – 30 分钟（持续优化）
数据利用率	仅基础统计	全维度 AI 标准化数据资产

3. 运营效率 – 成本指标

指标	改造前	改造后
专职巡检团队规模	20 人	12 人 + IoT AI 数据归拢
预测模型迭代周期	无，软件采购后模型冻结	依实际数据增量，可自选



AI 落地项目元数据

项目单期投入 8 – 12 周，具体根据数据规模和系统集成深度定制
(相关硬件与数据准备，需按实际情况另计)

FDE 交付团队 5 人（FDE Leader/Echo X 1、FDE 交付/Delta X 3、FDE PMO X 1）

商务联系

即刻与 采恪智创 客服联系，了解更多详细情况。

企微： 

电话： 0086-180-0179-2927

邮箱： siying.xu@usingnow.cn

