

利用小数据打造 AI 小模型：零售业低资源决策引擎开发

讲师 姚伟

【目标简述】

知识层面：

掌握小数据场景下特征工程与轻量化模型优化原理

理解零售核心指标（滞销率/周转率）的 AI 小模型驱动逻辑

熟悉端到端小模型解决方案框架（数据→建模→决策）

技能层面：

使用 Python 构建 零售小数据模拟器（生成/增强千级样本）

开发轻量级 AI 决策沙盘（含小模型预测+业务模拟模块）

部署低资源依赖型 AI 系统（万行数据以下场景）

实践层面：

用小模型解决零售业痛点：

小数据精准补货（动态安全库存算法）

小样本动态定价（价格弹性矩阵）

降低滞销率，提升周转率

【课程特色】

1. 小模型技术闭环

覆盖 轻量模型全链路：从数据合成（SMOTE）→特征优化→模型蒸馏（ONNX）→边缘部署（树莓派）

专为 低数据零售场景设计：适配便利店/商超 SKU 稀疏、样本不均衡问题

2. 零售沙盘实战驱动

开箱即用沙盘代码：

数据生成器：模拟 1000~5000 条零售样本

四步决策引擎：小数据训练→轻量模型预测→库存模拟→策略优化

业务规则深度耦合：动态定价代码直连价格弹性小模型

3. 小样本增强技术

数据层面：SMOTE-ENN 合成+外部变量融合（天气/竞品权重）

标签层面：Snorkel 弱监督标记（25%分位数滞销品自动标注）

模型层面：LightGBM 增量训练（周级 500 条数据热更新）

【课程简介】

本课程聚焦“小数据打造 AI 小模型”技术路径，以 零售业真实场景 为载体，深度融合：

小数据理论体系（稀疏数据价值 ROI、不均衡样本增强）

轻量化建模技术（LightGBM 增量训练、模型蒸馏部署）

零售决策引擎开发（动态定价/补货沙盘）

学员将通过 3 天实战：

1. 掌握小模型核心技术：解决零售业数据稀缺下的滞销预测、周转优化问题
2. 构建端到端解决方案：从千条数据生成→轻量模型训练→业务决策模拟
3. 获得跨行业迁移能力：复用方法论至制造业备件预测、医疗药品需求等小数据场景

【课程大纲】

课程名称	课程时间	课程大纲
小数据基础理论与零售场景建模	第一天 上午	<p>小数据驱动零售 AI 模型构建原理</p> <p>开营与行业洞察</p> <p>零售业小数据痛点与价值地图</p> <p>1.1 小数据科学核心</p> <p>小数据定义与特征（稀疏/不均衡/高噪声）</p> <p>与传统大数据对比矩阵</p> <p>小数据价值矩阵：ROI 公式与数据收集优先级决策树</p> <p>1.2 小数据小模型增强技术</p> <p>SMOTE-ENN 合成技术原理</p> <p>图神经网络关系建模：通过商品关联性提升预测精度</p> <p>外部变量融合体系：天气/节假日/竞品价格的动态权重配置</p> <p>1.3 零售核心指标建模</p> <p>滞销率=f(动销天数, 保质期, 价格弹性) 的动态计算模型</p>
		<p>库存周转率优化公式：AI 驱动的安全库存算法</p> <p>沙盘演示：价格变动对周转率影响的实时模拟</p>

	第一天 下午	<p>小数据特征工程与沙盘架构</p> <p>2.1 小数据特征工程</p> <p>动态特征构建框架：滚动周销量波动率 + 季节指数校正</p> <p>时序特征提取原理（周波动率/季节指数）</p> <p>特征交叉黄金组合：价格带×促销类型×时段的多维分析</p> <p>2.2 轻量化沙盘设计</p> <p>数据生成逻辑：参数化模拟器设计（SKU/店铺/季节的配置界面）</p> <p>业务规则嵌入设计：动态定价策略与自动补货规则的代码实现</p> <p>2.3 行业痛点解决方案</p> <p>便利店/商超小数据场景对比</p> <p>2.4 小数据标记体系设计（业务标签映射）</p> <p>滞销品动态标记规则：销量 < 历史 25 分位数的代码实现（关联沙盘 is_slow_mover 字段）</p> <p>高价值商品标签：价格弹性×周转率的复合公式</p>
小数据建模与轻量级 AI 开发实战	第二天 上午	<p>小数据模型构建技术</p> <p>1.1 轻量模型架构</p> <p>LightGBM 增量训练架构：每周新增 500 条数据的模型热更新</p> <p>迁移学习特征迁移方案</p> <p>1.2 可解释性工程</p> <p>SHAP 值原理：可视化价格变动对销量的边际影响</p> <p>实操 1：环境部署</p> <p>零售 AI 模拟沙盘基础环境部署</p> <p>1.3 弱监督标记技术</p> <p>小样本标记增强方案</p> <p>Snorkel 框架实战：用规则函数合成训练标签</p> <p>主动学习标记策略：基于模型置信度的样本筛选</p>
	第二天 下午	<p>端到端小数据小模型沙盘开发</p> <p>2.1 小数据小模型开发全流程</p> <p>特征工程流程：自动特征选择→异常值处理→编码→归一化</p> <p>列变换器配置技巧：类别特征嵌套编码的陷阱规避</p> <p>实操 2：沙盘操作</p> <p>从小数据生成到决策全流程实验（四步决策沙盘）</p> <p>小数据小模型四步决策实验：</p> <ol style="list-style-type: none"> 1. 生成小样本数据(5000 条) 2. 训练模型(R^2) 3. 参数预测 4. 库存模拟 <p>2.2 业务规则优化</p> <p>动态阈值算法：基于预测置信度的安全库存调整</p> <p>2.3 标记-建模联动优化</p> <p>动态标签迭代机制</p> <p>基于预测结果的标签自更新：预测滞销率 > 30% → 自动标记</p> <p>标签版本管理：Git 式标签溯源（防止标签漂移）</p>

小数据小模型行业迁移与轻量化部署	第三天上午	<p>跨行业小模型迁移策略</p> <p>1.1 零售沙盘洞察总结</p> <p>价格弹性矩阵：敏感商品与非敏感商品的定价策略差异</p> <p>最优库存周转公式：AI 建议 VS 传统经验对比实验</p> <p>1.2 小模型跨行业改造</p> <p>改造方法论：</p> <p>制造业备件预测：设备故障率×供应链时效的联合建模</p> <p>医疗药品需求预测：流行病学数据融合的小样本增强方案</p> <p>1.3 技术演进路径</p> <p>小数据解决方案演进路径：</p> <p>传统统计→机器学习→增强学习</p>
	第三天下午	<p>小模型部署与未来趋势</p> <p>2.1 轻量化企业部署</p> <p>轻量化部署架构：模型蒸馏+ONNX 运行时优化</p> <p>边缘计算方案：树莓派运行补货预测模型</p> <p>安全与合规要点：差分隐私在小数据中的应用</p> <p>2.2 前沿趋势</p> <p>联邦学习在小数据的应用</p> <p>生成式 AI 的数据增强前景</p>

【讲师简介】



姚伟

金融行业数据专家，国内 AI 5w 人社群私塾成员
某金融公司资深 DBA。拥有 10 余年数据库和大数据运维和管理经验。

深耕互联网行业 15 年以上，曾任数据库部门负责人。

作为国内 AI 5w 人社群私塾成员，长期聚焦全球 AI 前沿技术，尝试多种 AI 技术解决方案。历任资深 DBA、数据库专家、大数据负责人，参与各类数据系统与 AI 结合的各种项目，帮助企业 AI 提效和 AI 智能体相关解决方案。