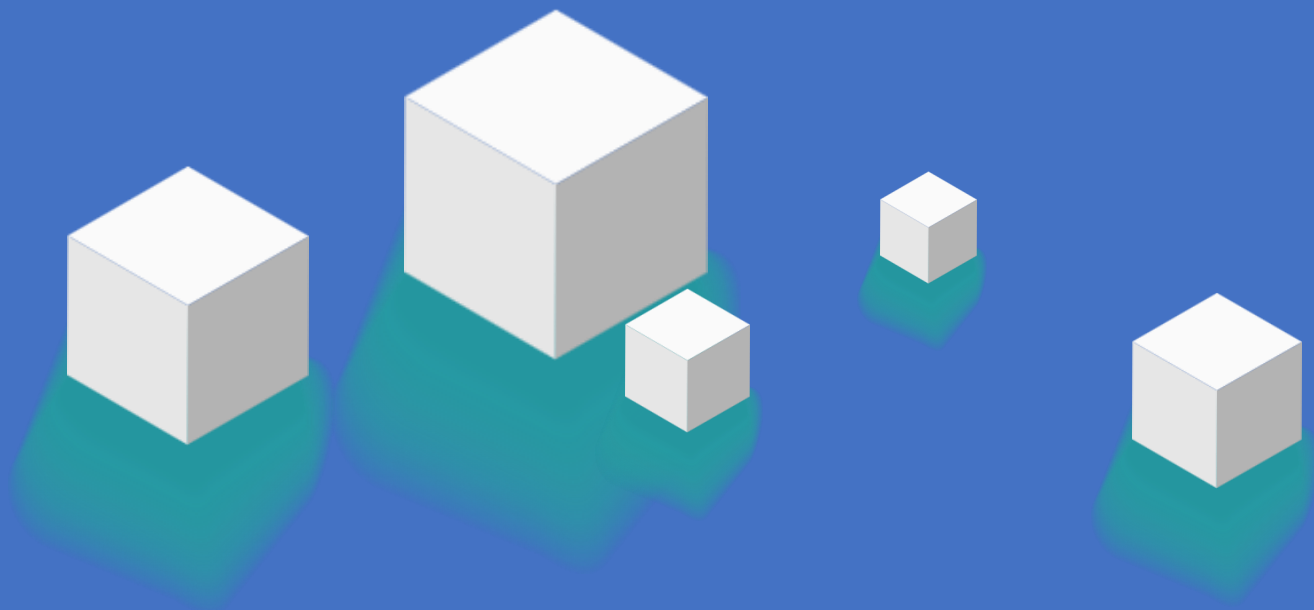


AI学习路径与规划



Self-introduction

和计算机、算法相关

(10岁，清华计算机博士，NOI，ACM比赛，ACM，IEEE，中国人工智能协会，阿里云MVP，CCF专委)

和培训、企业服务相关

(专栏付费订阅人数超过20万，企业客户包括：蚂蚁金服，美的，汽车之家，上汽大众，晶科能源，航天信息，中国银行，中国银联，花旗银行，汇丰银行，杭州银行，泰隆银行，中原银行，长沙银行，渤海银行，平安银行，平安保险，平安科技，阳光保险，兴业银行，兴业数金，上海银科，尚诚金融，嘉银金科，马上金融，中泰证券，富达基金，大恒集团，华润集团，中国外汇交易中心，联通软研院，梅赛德斯奔驰，雀巢，麦当劳，西门子等)



>> 今天的学习目标

AI学习路径与规划

- 课程安排
- 课程服务
- AI模块梳理
- 学习规划

课程安排

课程大纲

- 主干课

<https://rcne0lfduyr7.feishu.cn/wiki/VTBvwDKxJi1SPek52lXcfvrpn1w>



序号	课程小结	课程内容	学习收获
模块1: AI大模型应用开发基础			
1	AI大模型基本原理及API使用	<div>1. AIGC发展: 从GPT1到GPT4<ul style="list-style-type: none">- 什么是AI- 分析式AI与生成式AI- GPT是如何训练出来的- AIGC的表现与优势- AIGC的通用能力应用- CASE: 上手车险反欺诈 (基于ChatGLM)</div> <div>2. 大模型API使用<ul style="list-style-type: none">- 全球AI发展现状- CASE-情感分析-Qwen (掌握DashScope调用大模型)- CASE-天气Function-Qwen (了解Function Call)- CASE-表格提取-Qwen (了解多模态大模型)- CASE-运维事件处置-Qwen</div>	<div><ul style="list-style-type: none">- 理解AIGC与大模型的基本原理- 理解分析式AI与生成式AI的区别- 熟悉大模型API的调用方法- 学会使用DashScope等平台调用大模型 (如Qwen), 完成实际任务 (情感分析、表格提取等)</div>
2	DeepSeek使用与Prompt工程	<div>1. DeepSeek使用<ul style="list-style-type: none">- DeepSeek的创新- CASE: 小球碰撞试验 (Cursor + DeepSeek-R1)- DeepSeek私有化部署选择- Ollama部署DeepSeek-R1- API调用DeepSeek</div> <div>2. Prompt工程: 设计与优化<ul style="list-style-type: none">- Prompt原理- 提示词策略差异- 提示词关键原则- 提示词编写框架 (重要性排序)- 提示词编写技巧 (限制格式、区分、小样本学习、CoT、角色扮演等)</div> <div>3. Prompt工程实战<ul style="list-style-type: none">- CASE: 使用提示词完成任务- CASE: JSON格式返回- CASE: 使用CoT分步骤推理- CASE: 使用Prompt调优Prompt</div>	<div><ul style="list-style-type: none">- 了解DeepSeek的创新技术, 熟悉其API调用、本地部署 (Ollama) 及私有化方案- 理解Prompt工程的核心原理与优化策略- 学习提示词的设计逻辑、关键原则及主流框架 (如CoT、角色扮演等), 掌握不同场景下的Prompt优化技巧。- 通过案例实战 (如JSON格式生成、分步骤推理、Prompt自优化), 培养解决复杂问题的Prompt工程思维。</div>
3	Cursor编程-从入门到精通	<div>1. Cursor编程<ul style="list-style-type: none">- 什么是Cursor Rules- Cursor的主要功能</div> <div>2. Cursor编程实战<ul style="list-style-type: none">- CASE: 多张Excel报表处理- CASE: 疫情实时监控大屏</div> <div>3. Trae与CodeBuddy使用<ul style="list-style-type: none">- Trae使用- CodeBuddy使用</div>	<div><ul style="list-style-type: none">- 掌握 Cursor 编程的核心概念与功能 (如Cursor Rules, 熟悉 Cursor 的主要功能等)- 通过实际案例 (如多张 Excel 报表处理、疫情实时监控大屏) 掌握 Cursor编程- 掌握数据可视化能力, 使用Cursor可以搭建实时监控大屏- 了解 Trae 和 CodeBuddy 等其他AI工具</div>

序号	课程小结	课程内容	学习收获
模块2: AI大模型应用核心开发工具及技术			
4	Embeddings和向量数据库	<div>1. CASE：基于内容的推荐<ul style="list-style-type: none">- 什么是N-Gram- 余弦相似度计算- 为酒店建立内容推荐系统</div> <div>2. Word Embedding<ul style="list-style-type: none">- 什么是Embedding- Word2Vec进行词向量训练</div> <div>3. 什么是向量数据库<ul style="list-style-type: none">- FAISS, Milvus, Pinecone的特点- 向量数据库与传统数据库的对比</div> <div>4. Faiss工具使用<ul style="list-style-type: none">- Case: 文本抄袭自动检测分析- 使用DeepSeek + Faiss搭建本地知识库检索</div>	<div><ul style="list-style-type: none">- 掌握Embedding核心原理与应用- 理解N-Gram、词向量（Word2Vec）和余弦相似度等核心概念，学会训练和使用词嵌入模型- 了解掌握FAISS、Milvus、Pinecone等向量数据库的特点及适用场景- 通过酒店内容推荐系统、文本抄袭检测等案例，学会基于Embedding和向量数据库构建实际应用- 学习使用DeepSeek + FAISS构建本地知识库，实现高效语义检索与问答系统。</div>
5	RAG（Retrieval Augmented Generation）技术与应用	<div>1. 大模型开发的三种范式<ul style="list-style-type: none">- 提示词工程- RAG技术- 模型微调</div> <div>2. RAG技术<ul style="list-style-type: none">- 什么是RAG技术？它如何增强大模型的生成能力- RAG的核心原理与流程- NativeRAG- NoteBookLM使用</div> <div>3. CASE：DeepSeek + Faiss搭建本地知识库检索<ul style="list-style-type: none">- PDF文本提取与处理- 向量数据库构建- 语义搜索与问答链- 文本块对应页码信息源</div> <div>4. 如何提升RAG质量<ul style="list-style-type: none">- 数据准备阶段- 知识检索阶段- 答案生成阶段</div>	<div><ul style="list-style-type: none">- 掌握大模型开发的三大核心范式（提示词工程、RAG技术、模型微调）- 学习RAG的核心原理（检索-增强-生成），掌握从数据准备、向量检索到答案生成的完整技术链路。- 通过DeepSeek+FAISS案例，实现PDF文本处理、向量数据库构建及语义搜索问答链的开发。- 学习如何提升RAG系统的召回和回答能力</div>

序号	课程小结	课程内容	学习收获
6	RAG高级技术与最佳实践	<ol style="list-style-type: none">1. RAG技术树<ul style="list-style-type: none">- 模型不同阶段的RAG- RAFT方法：基于微调的RAG方法2. RAG高效召回方法<ul style="list-style-type: none">- 优化查询扩展- 双向改写- 索引扩展- Small-to-Big3. GraphRAG使用<ul style="list-style-type: none">- GraphRAG 过程- 全局搜索- 局部搜索- GraphRAG使用4. Qwen-Agent中的RAG<ul style="list-style-type: none">- 级别一：检索- 级别二：分块阅读- 级别三：逐步推理- RAG评测结果- Qwen-Agent使用	<ul style="list-style-type: none">- 了解RAG前沿技术体系（RAG技术树，包括模型阶段适配、RAFT微调方法等高级技术）- 掌握查询扩展、双向改写、索引扩展等高效召回方法，提升检索精度与覆盖率- 学习GraphRAG的全局/局部搜索机制- 使用Qwen-Agent的三级RAG实现（检索→分块→推理）
7	Text2SQL：自助式数据报表开发	<ol style="list-style-type: none">1. Text to SQL技术<ul style="list-style-type: none">- LLM模型选择- 大模型API使用- Function Call- 搭建SQL Copilot- LangChain中的SQL Agent- 自己编写（LLM + Prompt）2. 案例: 保险场景SQL Copilot实战	<ul style="list-style-type: none">- 学习大模型（LLM）在Text2SQL任务中的应用，掌握模型选型、API调用及Function Call等关键技术。- 通过LangChain SQL Agent或自研方案（LLM + Prompt），实现自然语言到SQL的自动转换，提升数据查询效率。- 结合保险行业需求，开发定制化SQL Copilot，完成从自然语言提问到精准SQL生成的端到端实现。

序号	课程小结	课程内容	学习收获
8	LangChain：多任务应用开发	<ol style="list-style-type: none">LangChain基本概念<ul style="list-style-type: none">- Models, Prompts, Memory, Indexes, Chains, Agents- LangChain中的tools (serpapi, llamamath)- LangChain中的Memory- LCEL构建任务链LangChain实战<ul style="list-style-type: none">- CASE：动手搭建本地知识智能客服（理解ReAct）- CASE：工具链组合设计（LangChain Agent）- CASE：搭建故障诊断Agent（LangChain Agent）- CASE：工具链组合设计（LCEL）AI Agent对比	<ul style="list-style-type: none">- 理解LangChain中的Models、Memory、Chains、Agents等核心组件- 掌握SerpAPI、LLM-Math等Tools的集成方法，学会用LCEL构建复杂任务工作流。- 通过知识客服、故障诊断等案例，实现具备记忆、推理和工具调用能力的智能体。- 对比不同Agent架构优劣，具备根据业务需求选择技术方案的能力。
9	Function Calling与智能Agent开发	<ol style="list-style-type: none">Function Calling的概念与应用<ul style="list-style-type: none">- 什么是Function Calling?- Function Calling 与 MCP的区别Qwen3 Function Calling使用<ul style="list-style-type: none">- Qwen3的特点- 使用Qwen3完成天气调用 Function Calling搭建业务助手<ul style="list-style-type: none">- Qwen-Agent中的 Function Calling- 使用Function Calling完成数据库查询- 数据表可视化	<ul style="list-style-type: none">- 理解Function Calling的概念、原理及其与MCP的区别- 学习Qwen3的特性，并通过实战案例（如天气查询）掌握其Function Calling的调用与集成方法。- 基于Qwen-Agent，利用Function Calling实现数据库查询、数据可视化等业务功能，完成端到端开发。
10	MCP与A2A的应用	<ol style="list-style-type: none">MCP的概念与应用<ul style="list-style-type: none">- 什么是MCP?- MCP 的核心概念 (MCP Host, MCP Client, MCP Server)- MCP 的使用场景- CASE：旅游攻略MCP- CASE：Fetch网页内容抓取- CASE：Bing中文搜索- CASE：桌面TXT统计器（MCP SDK使用）A2A的概念与应用<ul style="list-style-type: none">- 什么是Agent2Agent- A2A的关键组件- A2A的工作流程- A2A与MCP的关系- CASE：安排篮球活动（多智能体协作）	<ul style="list-style-type: none">- 理解MCP的核心概念（Host/Client/Server）及使用场景，- 通过旅游攻略生成、网页抓取、中文搜索、桌面TXT统计等案例，实现MCP的落地应用- 学习Agent2Agent（A2A）的工作流程与关键组件，掌握多智能体协同的任务分配与决策方法- 理解MCP与A2A的技术差异与互补关系

序号	课程小结	课程内容	学习收获
11	Agent智能体系统的设计与应用	<ol style="list-style-type: none">智能体的定义与作用<ul style="list-style-type: none">- 什么是AI Agent- AI Agent 的核心概念 (规划、记忆、工具调用等)- AI Agent框架对比 (LangChain, LangGraph, Qwen-Agent, Coze, Dify)- 什么时候用Agent- AI Agent与 workflows智能体分类<ul style="list-style-type: none">- 反应式 Reactive- 深思熟虑 Deliberative- 混合式 Hybrid智能体实战<ul style="list-style-type: none">- CASE: 私募基金运作指引问答助手- CASE: 智能投研助手- CASE: 投顾AI助手	<ul style="list-style-type: none">- 深入理解Agent的规划、记忆、工具调用等核心模块，掌握主流框架（LangChain/Qwen-Agent等）的技术差异与选型标准。- 学习反应式、深思熟虑式、混合式Agent的设计模式，能够根据业务需求选择最优架构。- 通过私募基金问答、智能投研、投顾助手等案例，完成不同类型的Agent实现。
12	视觉大模型与多模态理解	<ol style="list-style-type: none">VLM在行业中的应用<ul style="list-style-type: none">- Qwen-VL使用- Qwen-VL微调- 医疗行业中的应用（病历提取）- 车险承保中的应用（车辆承保、危险驾驶识别、车辆损失评估、事故要素提取、车辆一致性校验）视频理解SOTA<ul style="list-style-type: none">- InternVideo2/2.5- 视频基础模型- InternVideo2 预训练- 视频多模态注释示例- CASE: 汽车剐蹭视频理解MinerU使用<ul style="list-style-type: none">- PDF 转 Markdown- 网页内容提取- MinerU的核心技术- MinerU的应用场景- MinerU使用（在线使用、客户端）- MinerU使用（API使用）- MinerU私有化部署	<ul style="list-style-type: none">- 学习Qwen-VL等视觉大模型的微调与部署方法，掌握其在医疗、保险等领域的落地实践。- 理解InternVideo2/2.5等视频基础模型的预训练与多模态注释技术，实现场景下的视频分析（如汽车剐蹭）- 掌握MinerU的文档转换、内容提取及私有化部署能力，提升多模态数据处理效率

序号	课程小结	课程内容	学习收获
13	Fine-tuning微调艺术	<ol style="list-style-type: none">高效微调方法概述<ul style="list-style-type: none">- 参数高效微调 (Parameter-Efficient Fine-Tuning, PEFT)- 主要方法对比: Adapter, Prefix-tuning, LoRALoRA的数学原理<ul style="list-style-type: none">- 低秩分解与矩阵近似- 两个低秩矩阵的作用与优化微调数据准备<ul style="list-style-type: none">- 数据质量与数量要求- 不同模型尺寸与场景的数据需求硬件需求与显存计算<ul style="list-style-type: none">- 微调显存估算方法- LoRA显存优化与计算示例	<ul style="list-style-type: none">- 掌握高效微调技术 <p>学习LoRA、Adapter等方法，显著降低训练成本，提升大模型微调效率。</p> <ul style="list-style-type: none">- 深入理解LoRA数学原理 <p>掌握低秩矩阵分解的核心思想，灵活调整参数以适应不同任务需求。</p> <ul style="list-style-type: none">- 优化数据准备策略 <p>了解数据质量、数量对微调效果的影响，合理规划数据集规模。</p> <ul style="list-style-type: none">- 精准计算硬件需求 <p>学会估算显存占用，优化GPU资源，降低微调硬件门槛。</p>
14	Fine-tuning实操	<ol style="list-style-type: none">模型微调的方法<ul style="list-style-type: none">- 如何模型微调 (以ChatGLM为例)- 李飞飞 50美金 复刻R1模型- s1: Simple test-time scalingUnsloth: LLM高效微调<ul style="list-style-type: none">- CASE: qwen2.5-7B微调 (alpaca-cleaned)- CASE: 训练垂类模型 (中文医疗模型)- CASE: 训练自己的R1模型- CASE: YAML配置助手 (模型微调)- 打造金融垂类大模型 (智能客服)	<ul style="list-style-type: none">- 学习从基础微调 (如ChatGLM) 到高效优化技术 (如Unsloth)，掌握不同场景下的模型适配策略。- 掌握低成本微调技术，如何用50美金复刻自己的R1- 掌握中文医疗模型、金融客服模型等垂类大模型的训练与优化，解决行业特定需求

序号	课程小结	课程内容	学习收获
15	Coze工作原理与应用实例	<div>1. Coze工作原理<ul style="list-style-type: none">- Agent与Copilot的区别- 插件使用- 工作流使用- RAG知识库</div> <div>2. Coze应用实例<ul style="list-style-type: none">- CASE：AI新闻Agent- CASE：创建搜索新闻工作流- CASE：weather_news工作流（基于意图识别）- CASE：抖音文案提取&二创- CASE：LLM联网搜索- CASE：搭建古诗词Agent</div>	<div>- 理解Coze的核心机制：掌握Agent与Copilot的区别，熟悉插件、工作流及RAG知识库的应用逻辑。</div> <div>- 掌握Coze实战技能：通过案例学习，独立完成新闻Agent、天气新闻工作流等场景的搭建与优化。</div> <div>- 提升AI自动化效率：学会利用Coze实现联网搜索、文案提取与二创等任务，减少人工干预。</div>
16	Coze进阶实战与插件开发	<div>1. Coze进阶实战<ul style="list-style-type: none">- 批处理- Coze应用- 数据表使用- 多Agents模式- 多工作流复杂应用</div> <div>2. Coze应用实例<ul style="list-style-type: none">- CASE：古诗词绘画（批处理）- CASE：智能投顾助手（风险评测与推荐）- CASE：客户分层营销助手- CASE：智能客服Agent- CASE：市场舆情监测Agent</div>	<div>- 掌握Coze高阶功能应用：深入理解批处理、数据表操作、多Agents协作及复杂工作流设计，提升自动化任务处理能力。</div> <div>- 熟练开发Coze插件与集成：学习插件开发逻辑，实现与外部系统的数据交互，扩展Bot功能边界。</div> <div>- 实战复杂场景解决方案：通过金融、营销、客服等真实案例，掌握多行业场景的AI应用设计与优化技巧。</div>

序号	课程小结	课程内容	学习收获
17	Dify本地化部署和应用	<div>1. Dify本地化部署<ul style="list-style-type: none">- Dify开发平台- Docker Compose部署- 克隆Dify 代码仓库- 启动 Dify 服务- 访问 Dify- 如何使用 Dify</div> <div>2. Dify应用实战<ul style="list-style-type: none">- CASE：LLM联网搜索- CASE：搭建古诗词Workflow- CASE：智能客服ChatFlow- CASE：智能文档分析助手(MinerU+Dify)</div> <div>3. 如何应用Agent API（Coze, Dify）<ul style="list-style-type: none">- Coze API使用- Cozepy工具- Dify API使用</div>	<div>- 掌握Dify本地化部署：学习通过Docker Compose部署Dify，完成代码克隆、服务启动及平台访问全流程。</div> <div>- 熟悉Dify核心功能：理解Dify开发平台的操作逻辑，掌握Workflow、ChatFlow等工具的应用方法。</div> <div>- 实战AI应用开发：通过联网搜索、古诗词生成、智能客服等案例，掌握基于Dify的AI解决方案搭建。</div> <div>- 整合API扩展能力：学会调用Coze与Dify的API，实现跨平台自动化与功能扩展。</div>

序号	课程小结	课程内容	学习收获
18	分析式AI基础	<div>1. 分析式AI基础<ul style="list-style-type: none">- 分析式AI与生成式AI- 十大经典机器学习算法- 常用分类算法- 贝叶斯分类器- 决策树与随机森林- SVM支持向量机- 逻辑回归</div> <div>2. 项目实战<ul style="list-style-type: none">- CASE：二手车价格预测- 分类与回归的关系- 数据探索- 特征选择- 模型训练与预测</div>	<div>- 理解分析式AI核心概念：掌握分析式AI与生成式AI的区别，认识分析式AI在预测和决策中的独特价值。</div> <div>- 学习经典机器学习算法：深入理解十大经典机器学习算法，重点掌握分类、回归等核心算法的原理与应用场景。</div> <div>- 掌握数据建模全流程：从数据探索、特征选择到模型训练与预测，完整学习分析式AI项目的开发流程。</div> <div>- AI大赛实战能力：通过二手车价格预测等比赛实战，提升算法建模能力。</div>
19	不同领域的AI算法	<div>1. 金融行业的应用场景<ul style="list-style-type: none">- 银行不同部门的应用场景- 客户流失分析与预警- 因客定价- 长尾客群营销- 贷款商机挖掘- 商圈生意机会地图- 基于评分卡的风控模型开发- 期货套利模型- 资金流入流出预测</div> <div>2. 制造行业的应用场景<ul style="list-style-type: none">- 缺陷检测</div> <div>3. 快消行业的应用场景<ul style="list-style-type: none">- 供应链补货</div> <div>4. AI大赛：二手车价格预测（进阶）</div>	<div>- 掌握行业AI应用场景：理解AI在金融、制造、快消等不同领域的典型应用场景及业务价值。</div> <div>- 提升跨行业解决方案能力：通过缺陷检测、供应链优化等案例，培养将AI技术迁移至不同行业的能力。</div> <div>- 实战进阶：通过二手车价格预测大赛项目，强化特征工程的实战经验。</div>

序号	课程小结	课程内容	学习收获
20	机器学习神器	<div>1. 预测全家桶<ul style="list-style-type: none">- Project A: 员工离职预测- Project B: 男女声音识别- 分类算法: LR, SVM- 树模型: GBDT, XGBoost, LightGBM, CatBoost</div> <div>2. 机器学习神器<ul style="list-style-type: none">- 什么是集成学习- GBDT原理- XGBoost- LightGBM- CatBoost- AI大赛: 二手车价格预测- 如何防止模型过拟合</div>	<div>- 掌握核心机器学习算法: 系统学习分类算法 (LR、SVM) 和树模型 (GBDT、XGBoost、LightGBM、CatBoost) 的原理与应用。</div> <div>- 实战项目驱动学习: 通过员工离职预测、声音识别等案例, 掌握从数据预处理到模型训练的全流程。</div> <div>- 模型调优能力: 通过二手车价格预测大赛, 学习防止过拟合的策略, 提升模型泛化能力。</div>
21	时间序列模型	<div>1. 时间序列分析<ul style="list-style-type: none">- 什么是时间序列预测- statsmodels工具- AR、MA、ARMA、ARIMA模型</div> <div>2. 时间序列实战<ul style="list-style-type: none">- 使用ARMA/ARIMA对沪市指数进行预测- 资金流入流出预测</div>	<div>- 掌握时间序列分析基础概念: 理解时间序列预测的核心原理, 熟悉常用的统计模型 (AR、MA、ARMA、ARIMA) 。</div> <div>- 熟练使用分析工具: 学会运用Python的statsmodels库进行时间序列建模与预测。</div> <div>- 实战应用能力: 能够将ARMA/ARIMA模型应用于真实金融数据 (如沪市指数、资金流动预测) 。</div>
22	时间序列AI大赛	<div>1. Facebook Prophet分析<ul style="list-style-type: none">- 时间序列预测工具 prophet- 饱和增长- 突变点- 节日与大事件- Project A: 页面流量预测- Project B: 交通流量预测</div> <div>2. 时间序列AI大赛<ul style="list-style-type: none">- AI大赛: 资金流入流出预测- 使用prophet指数进行预测- 周期因子预测</div>	<div>- 掌握Prophet核心功能: 学习Facebook Prophet工具的核心特性, 包括饱和增长、突变点检测和节日效应建模。</div> <div>- 实战项目演练: 通过页面流量和交通流量预测项目, 掌握Prophet在真实场景中的应用。</div> <div>- AI竞赛能力提升: 参与资金流入流出预测大赛, 运用Prophet和周期因子优化预测效果。</div>

序号	课程小结	课程内容	学习收获
23	项目实战：企业知识库 (企业RAG大赛冠军项目)	<div>1. 企业RAG大赛：搭建RAG知识库<ul style="list-style-type: none">- RAG冠军方案（多路由+动态知识库）- RAG比赛任务说明- 基础RAG系统流程- 解析模块、Docling优化、表格序列化- 内容提取（ingestion）- 检索（Retrieval）- LLM 重排序 (LLM reranking)- 父页面检索- 整合后的检索器- 增强 (Augmentation)- 生成 (Generation)- 思维链、结构化输出、思维链+结构化输出- 指令细化 (Instruction Refinement)- 提示词创建、Prompt.py 实现- RAG系统调参</div> <div>2. 搭建自己的RAG系统<ul style="list-style-type: none">- 选择适合的LLM和Embedding模型- MinerU使用- 更新中文知识库、设置相关的问题清单- 针对开放式的问题，进行Prompt设置- 搭建前端页面，比如使用 streamlit</div>	<div>- 掌握RAG核心技术：深入理解RAG（检索增强生成）系统的工作原理，包括数据预处理、检索优化、生成增强等核心模块。</div> <div>- 实战冠军方案复现：学习RAG竞赛优胜方案，掌握多路由、动态知识库等高级技巧，并能优化基础RAG流程。</div> <div>- 搭建完整RAG系统：从零构建企业级知识库，涵盖模型选型、数据处理、Prompt优化到前端部署的全流程。</div>

项目实战

企业知识库

序号	课程小结	课程内容	学习收获
24	项目实战：交互式BI报表 (AI量化交易助手)	<div>1. ChatBI功能设计</div> <div>- 自然语言查询: 用户输入“查询2024年全年贵州茅台的收盘价走势”</div> <div>- 对比分析: 用户输入“对比2024年中芯国际和贵州茅台的涨跌幅”</div> <div>- 趋势分析: 用户输入“预测贵州茅台未来7天的收盘价”</div> <div>- 异常点分析 用户输入“检测贵州茅台近一年超买超卖点”</div> <div>- 新闻查询 用户输入“贵州茅台最近的热点新闻”</div> <div>2. 数据可视化与交互优化</div> <div>- 柱状图: 对比分析多个维度的表现</div> <div>- 折线图: 对历史数据进行趋势呈现</div> <div>- 交互设计: 支持用户自定义时间范围（如最近1个月、3个月、1年）</div> <div>3. 智能分析功能</div> <div>- 趋势预测: ARIMA模型（未来N天价格预测）</div> <div>- 异常检测: 布林带分析（超买/超卖点识别）</div> <div>- 周期性分析: Prophet模型（趋势分解）</div> <div>4. 知识库管理</div> <div>- few shot: 撰写few shot示例，提供大模型处理逻辑的准确性</div>	<div>- 掌握交互式BI系统开发：学习如何构建支持自然语言查询的BI系统，实现从数据采集到可视化分析的全流程开发。</div> <div>- 熟练应用智能分析模型：掌握ARIMA、布林带、Prophet等模型在智能分析中的实际应用。</div> <div>- 优化数据交互体验：设计高效的数据查询与可视化方案，提升用户与系统的交互体验。</div> <div>- 提升问题解决能力：通过实战项目，培养从需求分析到技术实现的完整问题解决能力。</div>

项目实战

交互式BI报表

序号	课程小结	课程内容	学习收获
25	项目实战：AI智慧运营助手 (百万客群经营)	<div>1. 智能客户洞察系统</div> <div>- 多模态客户画像</div> <div>- 动态标签生成：利用大模型生成客户动态标签（如“高净值”，“潜在高净值”等）。</div> <div>- 智能分析细化：基于智能分析模型（分类、聚类等）进一步细化客户画像，为后续精准营销提供更丰富的客户特征维度。</div> <div>2. 企业知识库引擎</div> <div>- RAG增强知识库：营销话术库构建</div> <div>- 大模型解释：利用SHAP分析，对个体客户的经营决策进行解释，增强客户经理执行的说服力</div> <div>3. 智能营销Agent（多Agent协作架构）</div> <div>- 分析Agent：价值判断，调用决策模型（如逻辑回归、决策树、LightGBM等）判断客户价值。</div> <div>- 推荐Agent：个性化方案生成，结合知识库（包括SHAP分析结果等）生成个性化方案</div> <div>- 话术Agent：沟通脚本生成，实时生成沟通脚本，融入关联分析（Apriori）得出的产品组合推荐话术。</div> <div>4. 可视化大屏搭建与对话式BI系统</div> <div>- 可视化大屏搭建：展示百万经营关键指标</div> <div>- 对话式BI系统：客户经理询问具体个体客户信息，通过多Agent协作，完成分析和推荐</div>	<div>- 掌握AI驱动的客户洞察技术：学习多模态客户画像构建与动态标签生成，利用大模型和智能分析模型（分类、聚类等）细化客户特征，提升精准营销能力。</div> <div>- 构建企业级知识库与决策解释系统：通过RAG增强知识库，结合SHAP分析，实现营销话术优化与客户经营决策的可解释性，增强执行说服力。</div> <div>- 设计智能营销Agent协作架构：掌握多Agent（分析、推荐、话术）协同 workflow，结合决策模型（如LightGBM）和关联分析（Apriori），实现自动化客户价值判断与个性化方案生成。</div> <div>- 实现可视化大屏与智能AI交互：搭建可视化大屏监控关键指标，开发对话式AI系统，支持客户经理通过自然语言交互快速获取客户洞察与营销建议。</div> <div>所用技术：RAG、Function Calling、向量数据库、Prompt工程、Agent技术运用</div>

项目实战

AI运营助手

序号	课程小结	课程内容	学习收获
26	项目实战：AI搜索类应用 (知乎直答)	<div>1. AI搜索类应用的架构设计：<ul style="list-style-type: none">- 数据准备与索引构建- 搜索模块开发（基于关键词的检索/基于语义的检索/混合检索）- 大模型推理和回复- 前端开发与用户体验</div> <div>2. AI搜索类应用开发<ul style="list-style-type: none">- Version1：对于多文件快速进行检索和回答- Version2：海量文件快速索引（ES）- Version3：添加向量检索功能- Version4：添加外部数据源（AI搜索MCP）- Version5：界面美化</div> <div>3. 基于Qwen-Agent的最佳实践<ul style="list-style-type: none">- CASE：长对话检索与问答- CASE：多文档并行问答- CASE：多智能体问答- CASE：多智能体协作路由</div>	<div>- 掌握AI搜索架构：学习混合检索系统（关键词+向量）的设计与实现，包括数据索引、搜索优化和大模型集成。</div> <div>- 完成版本迭代开发：从单文件检索逐步升级到海量数据+外部数据源的完整AI搜索应用。</div> <div>- 熟练应用Qwen-Agent：实现复杂场景下的智能问答（长对话、多文档、多智能体协作）。</div> <div>- 提升工程与体验优化能力：结合大模型推理与前端交互，打造高效、用户友好的搜索产品。</div> <div>所用技术:ElasticSearch、向量数据库、Prompt工程、RAG术、Qwen-Agent最佳实践</div>

项目实战

AI搜索类应用
(知乎直答)

<https://zhida.zhihu.com/>

【赠课】

课节	主要内容	课程目标
预训练+微调的训练范式 开源生态和OpenAI的差异详解	<div><div>1. 预训练阶段和SFT阶段并没有拉开差距</div><div>2. RLHF为何追不上OpenAI<div><div>a. 首席科学家Ilya Sutskever 的主要成就</div><div>b. Instruct GPT的论文中公开了哪些RLHF的具体细节</div><div>c. 尝试Reward Model时的常见问题</div><div>d. Proximal Policy Optimization的实操复杂度</div></div></div><div>3. 开源生态的另辟蹊径<div><div>a. Direct Preference Optimization的重要尝试</div><div>b. Reinforced Token Optimization的关键进步</div></div></div><div>4. 开源生态如何成功追赶OpenAI</div></div>	<div><div>1. 了解开源生态发展史，深入解读开源模型复现GPT进程其效果差距</div><div>2. 掌握不同开源模型与闭源模型技术差异及其原因</div><div>3. 具备项目负责人对模型选型的基础能力</div></div>
探索神经网络的奥秘	<div><div>1. 设计数学模型的主要工作是设计数学公式<div><div>a. 一个公式识别鳄鱼与蛇</div><div>b. $y=ax+b$ 就是一条直线</div><div>c. 参数a控制直线旋转，参数b控制直线平移</div><div>d. 人类数学家如何确定a和b的值</div></div></div><div>2. 设计数学公式的人类极限</div><div>3. 万金油公式 - 神经网络<div><div>a. 神经网络经典定义</div><div>b. 模拟人脑神经元的电信号传导</div></div></div><div>4. 案例：基于MNIST数据库的图像识别<div><div>a. 预处理图片数据</div><div>b. 设定神经网络公式结构</div><div>c. 设计神经网络参数提取特征数据</div><div>d. 多层网络提取深度特征</div></div></div><div>5. 神经网络与机器学习<div><div>a. 损失函数评定当前模型水平</div><div>b. 梯度下降确定模型迭代方向</div><div>c. 一次迭代提升模型精准程度</div></div></div></div>	<div><div>1. 打开黑盒，让专属于算法科学家们的神秘模型，用最通俗易懂的方式展现在众人面前</div><div>2. 深入理解方案专家必备的原理知识，设计方案时游刃有余，也让产品和应用方向更具长期性</div><div>3. 理解算法工程师、算法科学家到底在做什么样的工作，跨部门多角色协作奠定基础</div></div>

课程服务

课程服务

- 学习社群：讲师+助教+班主任
 - 任何专业问题都可以在群内提问
 - 学习过程中，任何问题可以找班班处理，如果有涉及隐私或者保密的专业性问题，不想在群内提问，也可以找班班转达提问

课程服务

- 线上环境、算力、课程开通等
 - 随着课程进程，班班会提前安排同学们开通，开通过程中有任何问题，可以私聊班班进行处理。

AI模块梳理

AI模块梳理

1. DeepSeek系列

- DeepSeek-V3、DeepSeek-R1等模型的推理部署（蒸馏、量化优化）
- API调用及性能调优。

2. 上手生成式AI

- Prompt Engineering（提示工程）：基础技巧、高级方法（包括CoT, Few-Shot等）
- Cursor/Trae应用：完成数据可视化、仿真建模、特征洞察等一系列任务，开发AI大模型的应用

掌握Prompt Engineering与Cursor使用，是大家后续工作和学习的基础

AI模块梳理

3. Qwen系列

- Qwen全系模型（72B/7B/Turbo/VL）的使用
- 多模态（Qwen-VL）开发
- 智能体（Qwen-Agent）开发，使用Qwen-Agent的自动化任务编排能力。
- 推理模型 QwQ 的应用
- 垂直领域的微调

Qwen系列依然是最好用的生成式AI，生态丰富，包括了文本、视觉、智能体、推理等一系列的模型，并且基座模型优秀，不断更新发展中

AI模块梳理

4. AI Agent系列

- RAG：结合检索与生成的问答系统搭建，重点解决外部知识库集成问题。
- Text2SQL：自然语言转数据库查询，适合企业数据交互场景。
- MCP开发：Model Context Protocol，模型上下文协议应用与开发
- LangChain应用：构建LLM应用的开发框架，提供链（Chains）、代理（Agents）、记忆（Memory）等核心组件，支持连接外部数据源和工具，实现问答系统、文本生成等复杂功能
- 向量数据库应用：Embedding训练与使用，Faiss应用
- Coze：工作流编排，插件开发
- Dify：基于LLM构建自主Agent，函数调用与工具链

AI Agent开发将是大家未来做项目中的主要工具，用于完整的AI项目的呈现

AI模块梳理

5. 机器学习系列

经典算法（线性回归、决策树、SVM、XGBoost、LightGBM等）的实战，覆盖特征工程、模型选择与调参。

机器学习在企业中有非常多的应用场景，包括：仿真建模、趋势预测、分类决策等

6. AI大赛实战

以天池“资金流入流出预测”为例，完整复现数据清洗、特征构建、时序建模（如ARIMA、Prophet）、模型融合等流程。

分析式AI与企业决策息息相关，将是AI应用中的重要一环，并通过生成式AI进行Chat互动，给用户提供友好的体验和使用

学习规划

学习规划

Thinking: AI可以写代码, 并且完成的比一般人好; AI可以画图, 并且质量比普通设计师强; 未来的工作实现, 不一定是人来完成具体的实现, 而是通过AI工具一起来进行。那么人的主要能力是什么?

- 承接需求: 通过各种方式, 完成任务, 任务驱动、需求驱动
- 问题定义: 识别真正需要解决的问题 (例如区分“用户要什么”和“用户需要什么”)。
- 广泛的知识面: 细节可以通过AI来互动和执行, 但是人需要广泛的知识面
- 质量检查: AI可能会出错, 比如参考文献、时间等, 看起来通顺的内容, 不一定真实存在
- 深度思考: 将人从碎片化的时间中解放出来, 深度思考 (没有打扰的情况下, Think Big, not small)
- 成果展示: 向上管理, 呈现自己的作品和交流

新的时代来临, 人将逐渐分化。有些人的需求会越来越多, 有些人的需求会越来越少 => 什么是关键?

学习规划

Thinking: 如何学好课程的内容?

- Learning 整理笔记: 将关键内容进行梳理, AI无法替代你的思考和整理
- Challenge 问题驱动: 思考你工作中遇到的重要的问题, 有价值/共性的问题, 并将它具象化
- Show 作品呈现: 和AI一起工作, 不断完成作品, 与同学们进行分享

Thinking: 小白如何学好AI?

要花更多的时间 (practice) + 更聪明的学习 (filter)

- 实践: 实践是学习的核心部分, 通过实际操作, 可以更好地巩固知识 => 要花更多的时间
- 交流: 与他人交流、提问并接受反馈、获取新的见解 => 快速提升, 突破升级
- 培训: 包括课堂学习等正式学习 => 提供系统性的知识, 明确学习方向







学习规划

Thinking: 对工作有什么帮助?

AI应用是各企业的增长和盈利的最终导向，也是企业未来的资金投入导向

技术驱动：懂技术，掌握AI大模型技术，能够胜任的岗位：AI应用开发工程师、agent开发工程师、AI测试、AI运维等AI开发、测试、运维等AI技术类相关岗位

业务驱动：懂业务，掌握AI大模型技术，能够胜任的岗位：AI产品经理、AI项目经理、AI解决方案专家等业务驱动型AI相关岗位



Thank You
Using data to solve problems