| 阶段 | 时间范围 | 关键任务 | 交付成果 |

|---------------|----------------|------------------------------|--------------------------|

| Ⅰ 选题定位 | 第1-2周 | 确定创新点、文献缺口 | 研究问题框架图 |

| Ⅱ 实验设计 | 第3-4周 | 方法论验证、数据采集方案 | 实验流程图+时间预估表 |

| Ⅲ 数据攻坚 | 第5-8周 | 核心数据产出、可视化处理 | 数据图表包（含原始数据） |

| Ⅳ 初稿撰写 | 第9-11周 | 结构化写作、初版结论提炼 | 完整论文草稿（含附录） |

| Ⅴ 润色优化 | 第12-13周 | 同行预审、语言润色、格式精修 | 终稿PDF+投稿清单 |

| Ⅵ 精准投稿 | 第14-16周 | 期刊匹配、附信撰写、系统提交 | 投稿回执+跟进计划表 |

工作技能：

**四个月学习规划（2025年02月22日-2025年06月22日）‌**

基于推荐算法岗位核心要求与搜索结果的效率优化策略‌12，分阶段制定以下计划：

**‌第1个月：编程与算法基础强化‌**

‌**目标**‌：掌握Python/C++基础语法，刷题150+（高频题型为主），完成一个数据处理小项目。  
‌**每周安排**‌：

* ‌**周一至周五（每日3h）**‌：
  1. ‌**算法刷题（1.5h）**‌：完成3道LeetCode题（优先[字节高频题库](https://codetop.cc/)），重点突破动态规划（背包问题）、树结构（DFS/BFS）‌16。
  2. ‌**编程语言（1h）**‌：
     + Python：学习Pandas数据处理（多表Join、特征工程）‌14。
     + C++：掌握STL容器、内存管理（《Effective C++》第1-4章）‌46。
  3. ‌**复盘（0.5h）**‌：整理错题笔记，标注时间/空间复杂度优化思路‌5。
* ‌**周末（3h）**‌：
  1. 用Python实现用户行为日志分析工具（数据清洗+统计报表）‌24。

‌**关键产出**‌：

* GitHub仓库：上传LeetCode题解（含优化思路）和数据处理项目代码‌46。
* 掌握高频题型：动态规划、树操作、字符串处理‌16。

**‌第2个月：机器学习与推荐系统入门‌**

‌**目标**‌：掌握经典推荐模型（Wide&Deep、DIN），复现1篇顶会论文，Kaggle竞赛实战。  
‌**每周安排**‌：

* ‌**周一至周三（每日3h）**‌：
  1. ‌**理论学习（1h）**‌：精读《百面机器学习》第1-4章，整理LR/GBDT/DNN公式推导‌12。
  2. ‌**模型复现（1.5h）**‌：用PyTorch实现Wide&Deep模型（数据集：MovieLens），优化CTR指标‌24。
  3. ‌**论文精读（0.5h）**‌：阅读《Deep Interest Network》，标注用户兴趣建模方法‌23。
* ‌**周四至周五（每日3h）**‌：
  1. ‌**召回系统（1h）**‌：学习双塔模型与Faiss向量检索‌24。
  2. ‌**Kaggle竞赛（1.5h）**‌：参与[H&M个性化推荐赛](https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations)，优化Embedding生成策略‌24。
  3. ‌**复盘（0.5h）**‌：记录模型调参过程与效果对比‌5。
* ‌**周末（3h）**‌：复现DIEN模型核心模块（用户序列兴趣提取）‌23。

‌**关键产出**‌：

* Kaggle竞赛排名：Top 25%‌24。
* GitHub项目：复现Wide&Deep和DIEN模型‌24。

**‌第3个月：推荐系统全链路与业务适配‌**

‌**目标**‌：搭建召回→排序→重排全链路系统，掌握实时特征工程（Flink），研究电商推荐场景。  
‌**每周安排**‌：

* ‌**周一至周二（每日3h）**‌：
  1. ‌**召回模块（1h）**‌：搭建双塔模型+Faiss索引（优化ANN检索效率）‌24。
  2. ‌**排序模型（1.5h）**‌：实现多目标排序（MMOE模型，平衡CTR/CVR）‌24。
  3. ‌**复盘（0.5h）**‌：对比不同召回策略的Recall@20指标‌5。
* ‌**周三至周四（每日3h）**‌：
  1. ‌**重排策略（1h）**‌：设计多样性控制算法（DPP算法）‌24。
  2. ‌**实时特征（1.5h）**‌：学习Flink处理用户点击流数据（窗口函数、状态管理）‌24。
  3. ‌**业务分析（0.5h）**‌：研究字节电商推荐场景（冷启动、多样性优化）‌34。
* ‌**周末（3h）**‌：全链路联调，输出性能报告（QPS≥5k，Recall@20提升15%）‌24。

‌**关键产出**‌：

* 项目成果：完整推荐系统（含实时特征与多目标排序）‌24。
* 技术文档：系统设计流程图与性能优化方案‌25。

**‌第4个月：项目优化与求职冲刺‌**

‌**目标**‌：完善简历与作品集，模拟面试，补充AIGC+推荐前沿技术。  
‌**每周安排**‌：

* ‌**周一至周二（每日3h）**‌：
  1. ‌**简历优化（1h）**‌：量化项目成果（如“CTR提升12%”），突出Kaggle竞赛与论文复现‌12。
  2. ‌**作品集（1.5h）**‌：整理GitHub项目文档（含单元测试与性能报告）‌24。
  3. ‌**前沿技术（0.5h）**‌：学习生成式推荐（Diffusion Model+用户画像）‌34。
* ‌**周三至周四（每日3h）**‌：
  1. ‌**模拟面试（2h）**‌：高频题型训练（动态规划、推荐算法调优）‌16。
  2. ‌**笔试真题（1h）**‌：刷[牛客网字节历年题](https://www.nowcoder.com/)‌56。
* ‌**周末（3h）**‌：技术方向复盘（查漏C++内存管理、多线程编程）‌46。

‌**关键产出**‌：

* 简历亮点：Kaggle Top 25%、复现3篇顶会模型‌24。
* 面试题库：整理50+高频题解析（含代码与逻辑推导）‌56。

**‌每日时间表示例‌**

| **‌时间段‌** | **‌内容‌** |
| --- | --- |
| ‌**19:00-20:30**‌ | 算法刷题（LeetCode高频题+错题复盘）‌16 |
| ‌**20:30-21:30**‌ | 理论学习（机器学习模型/论文精读）‌23 |
| ‌**21:30-22:00**‌ | 项目开发（推荐系统模块实现）‌24 |

**‌效率优化建议‌**

1. ‌**优先级管理**‌：前两个月侧重算法与模型，后两个月聚焦系统设计与面试‌25。
2. ‌**量化反馈**‌：每周统计LeetCode通过率、Kaggle排名，动态调整计划‌15。
3. ‌**资源复用**‌：直接使用Kaggle数据集与开源项目（如微软Recommenders）加速开发‌24。

通过此方案，4个月可系统覆盖岗位要求的90%技能点，显著提升竞争力‌24。