

学习指南

夏晓波 中国科学技术大学

2026/3/19

学习目标

夯实机器学习与多模态学习理论基础，强化逻辑分析能力与创新思维能力。

学习内容

以下学习内容的理解深度与消化节奏存在个体差异。整体建议在 6-9 个月内完成一次系统性覆盖，无需对每一部分都做到完全理解，但应达到在今后接触相关概念时，能够迅速建立基本认知并判断其研究背景与使用场景。

1. 合理使用大语言模型（Gemini/ GPT/ Qwen/Doubao）作为学习辅助与信息检索工具：
 - 参考：Google 等平台上的相关科普材料及大语言模型使用指南、技术博客。
 - 目标：培养利用大语言模型获取信息、进行知识归纳整理与辅助问题求解的能力。

以下内容由浅入深

2. Python Programming (Deep learning 实现基础)
 - 参考：Google 上很多类似资源
 - 目标：熟练用 Python 编写常见机器学习算法（推荐 PyTorch）
3. 深度模型基本架构
 - 参考①：书籍《Deep Learning》；地址：<https://www.deeplearningbook.org/>；作者：Ian Goodfellow, Yoshua Bengio, and Aaron Courville；重点：“Part II: Modern Practical Deep Networks”。

- 参考②：斯坦福 CS336 课程；主讲人：Prof. Tatsunori Hashimoto & Prof. Percy Liang；课程地网页及视频：<https://stanford-cs336.github.io/spring2025/>；https://www.youtube.com/playlist?list=PLoROMvodv4r0Y23Y0BoGoBGgQ1zmU_MT
- 目标：理解不同网络架构（特别是大语言模型结构），熟练网络训练算法。

4. 机器学习基本概念

- 参考①：书籍《The Elements of Statistical Learning》；地址：<https://hastie.su.domains/Papers/ESLIII.pdf>；作者：Jerome H. Friedman, Robert Tibshirani, and Trevor Hastie；重点：Chapter-2, 3, 4, 12, 13, 16。
- 参考②：书籍《Foundations of Machine Learning》；地址：<https://cs.nyu.edu/~mohri/mlbook/>；作者：Mehryar Mohri；重点：Chapter-1, 2, 3, 4, 5。
- 目标：深入理解机器学习核心概念，了解其核心概念之间数学原理，建立知识框架。

5. 多模态机器学习基础知识

- 参考①：综述《Foundations and Trends in Multimodal Machine Learning: Principles, Challenges, and Open Questions》；地址：<https://arxiv.org/abs/2209.03430>；作者：Paul Pu Liang, Amir Zadeh, and Louis-Philippe Morency。
- 参考②：《MML Tutorial》；地址：<https://cmu-multicomp-lab.github.io/mml-tutorial/schedule/>。
- 目标：理解多模态机器学习基础知识与理论，以及其相较于单模态机器学习的特殊性和理论分析难点。

6. 多模态大模型基础知识

- 参考①：斯坦福 CS231N Slides；地址：https://cs231n.stanford.edu/slides/2025/lecture_16.pdf
- 参考②：斯坦福 CS25: V4|From Large Language Models to Large Multimodal Models；地址：<https://www.youtube.com/watch?v=cYfKQ6YG9Qo>
- 目标：理解常见多模态大模型的组件和关键技术，以及其相较于语言大模型的特殊性和构建技术难点。

应避免的学习误区

1. 不应以“通读教材”为目标，刻意追求线性、完整的阅读过程：
 - 科研与项目学习并非一次性投入、一次性产出的过程，应在不同阶段反复回看、间断式研读相关书籍与资料。
 - 随着研究背景与实践经验的积累，对同一内容理解会不断加深，这是正常且必要的过程。
2. 不应以“应试记忆”为导向，片面追求知识点的全面覆盖与精确复现：
 - 逻辑思维能力来源于对不同方法、模型与技术路线的对比分析。
 - 创新性思维能力来源于对不同概念、思想与视角的交叉理解与融合。
3. 不应将科研等同于课堂教学，期待现成、完整的提纲与清晰路线：
 - 对于前沿问题或创新性研究方向，通常不存在标准化、手把手的学习路径。
 - 学习者需要借助书籍与资料形成自身的基本认知框架与判断原则，并结合科研实践与项目经验，逐步构建和完善个人知识体系。
4. 不应将大语言模型视为“标准答案”，对其输出结果不加甄别地接受：
 - 大语言模型可能生成看似合理但实际并不准确的信息。
 - 在使用过程中，应始终保持批判性思维，通过查证、交叉比对与独立思考来验证和吸收相关内容。