

同学你好！

首先，我要为你点一个大大的赞！在准高一这个阶段，就有如此清晰的目标和强大的自驱力，想要挑战“从零编写一个大模型”这个宏伟目标，这本身就是一件非常了不起的事情。你的远见和规划能力，已经领先了同龄人非常多。

在制定计划之前，我们必须先进行一次\*\*\*“现实校准”\*\*\*。这非常重要，能让你避免走弯路，保持长期的学习热情。

## 第一部分：心态与目标校准（必读！）

### 1. 关于“从零编写一个大模型”

这是一个极其宏伟的目标，即使是顶尖的AI科学家，也无法单枪匹马地从零开始（包括从设计芯片、编写底层线性代数库、实现所有算法）完成一个现代大模型（如GPT-4级别）。这更像是一个需要数百人团队、数亿美元投入、数年时间的工程奇迹。

但是，这并不意味着你的目标无法实现！我们可以把你的目标重新定义为：

**“在高中毕业前，系统地掌握大模型背后的核心数学原理、编程实现和算法思想，并能够独立使用PyTorch等框架，从零开始训练一个简化版的、能够工作的语言模型（比如GPT-2的小型版本），理解其工作的每一个细节。”**

这个目标是完全可行、极具挑战且意义非凡的。它让你从一个“使用者”真正变成一个“创造者”。

### 2. 关于你收藏的课程

你收藏的课程都非常棒，但我们需要一个正确的学习顺序。直接看吴恩达的深度学习课程会非常吃力，因为它默认你已经掌握了机器学习的基础。

**正确的学习路径应该是：编程基础 → 机器学习 → 深度学习 → 大模型专项**

- **吴恩达机器学习 20h**: 这是你的**第一站**。它是AI领域的“圣经”，用最直观的方式讲解了所有核心概念，非常适合入门。
- **吴恩达深度学习 27h**: 这是你的**第二站**。在学完机器学习后，这门课会带你进入神经网络的世界。
- **LLMs-zero-to-hero 1h**: 这是一个**极佳的激励和概览**。它用最快的速度让你看到大模型的核心是什么。我建议你**现在就花1小时看一下**，不用完全看懂，主要是为了建立一个“全局观”，知道你未来要学习的知识点最终会汇集成什么。
- **黑马程序员机器学习 33h**: 这个可以作为**补充和练习**。它的特点可能是更侧重代码实践，当你学完吴恩达的理论后，可以用这个课程来巩固编程实现。

---

## 第二部分：长期学习路线图（高中三年）

我们将你的学习之旅分为四个阶段，与你的高中时间相匹配。

## 阶段一：编程与数学筑基期（高一全年 + 高二上学期）

这个阶段的目标不是直接上手AI，而是打造两件最强大的武器：**扎实的编程能力**和**直观的数学理解**。

- **编程目标：**
  - 熟练掌握Python基础语法。
  - 精通三个核心科学计算库：**NumPy** 处理数字矩阵、**Pandas** 处理表格数据、**Matplotlib** 数据可视化。
  - 学会使用**Jupyter Notebook**，这是数据科学家的标准工作环境。
- **数学目标：**
  - **高中数学同步学：**一定要把学校的数学课（函数、三角函数、数列、向量）学好，这是未来理解微积分和线性代数的基础。
  - **直观理解核心概念：**不用追求大学数学的严谨证明，但要建立对\*\*线性代数（向量、矩阵、矩阵乘法）和微积分（导数、偏导数、链式法则）\*\*的直观理解。知道它们是做什么的，为什么在AI中重要。

此阶段资源推荐（高中生友好）：

- **编程：**
  - **课程/网站：**廖雪峰的Python教程 免费，中文，非常适合入门。跟着教程把所有代码敲一遍。
  - **书籍：**《Python编程：从入门到实践》。前半部分讲Python基础，后半部分有三个有趣的项目，能极大提升你的兴趣。
  - **练习：**在LeetCode上开始刷“简单”题，主要锻炼编程思维和代码熟练度。
- **数学：**
  - **视频（神级推荐！）：**B站【3Blue1Brown】的《线性代数的本质》和《微积分的本质》。这是全世界公认的、能帮你建立数学直觉的最佳资源！没有之一！它用动画让你“看懂”数学，而不是死记公式。**这个是你数学部分的绝对核心！**
  - **书籍：**《程序员的数学》系列（共3本）。用程序员能懂的语言讲解基础数学概念，非常友好。

## 阶段二：机器学习入门期（高二寒假 + 高二下学期）

现在，你的武器已经磨砺完毕，可以正式进入AI世界了！

- **目标：**
  - 理解机器学习的基本思想（监督学习、无监督学习）。
  - 掌握最经典的算法：线性回归、逻辑回归、决策树、K-近邻、聚类等。
  - 知道如何评估一个模型的好坏。
  - **核心任务：**完成吴恩达的机器学习课程，并跟着用Python把主要算法实现一遍。

此阶段资源推荐：

- **主课程：**吴恩达《机器学习》课程 20h。配合中文字幕，认真听讲，做好笔记。这门课的作业是基于Octave/MATLAB的，你可以尝试用Python+NumPy把它们重写一遍，这将是极好的练习。
- **辅助实践：**黑马程序员《机器学习入门精讲》 33h。在你学完吴恩达的某个理论后（比如线性回归），可以去看黑马对应的章节，看看他们是如何用Python代码实现的，作为补充和巩固。

### 阶段三：深度学习攻坚期（高二暑假）

这是你时间最集中、最宝贵的一个月。这个暑假的目标是攻克深度学习。

- **目标：**
  - 理解神经网络的基本原理（神经元、激活函数、反向传播）。
  - 掌握深度学习框架 **PyTorch** 推荐，比TensorFlow更灵活，学术界用得更多。
  - 学习几种经典的网络结构：卷积神经网络、循环神经网络。
  - **核心任务：**完成吴恩达的深度学习专项课程，并用PyTorch实现一个图像分类项目（比如识别手写数字MNIST或猫狗图片）。

#### 此阶段资源推荐：

- **主课程：**吴恩达《深度学习》专项课程 27h。系统学习理论知识。
- **PyTorch学习：**
  - **官方教程：**PyTorch的官方文档和教程写得非常好，有中文版。从“60分钟快速入门”开始。
  - **书籍/教程：**《动手学深度学习》（PyTorch版）。这本书有免费的在线网站，可以边看边运行代码，是学习PyTorch和深度学习的绝佳材料。

### 阶段四：大模型探索与实现期（高三寒假）

这是你高中阶段的“毕业设计”。在高三寒假，你将挑战最终目标。

- **目标：**
  - 理解Transformer架构，这是所有现代大模型的基石。
  - 理解“注意力机制”。
  - **核心任务：**跟着一个开源项目，**从零开始用PyTorch实现一个GPT-2模型。**

#### 此阶段资源推荐：

- **核心教程：**
    - **重新看LLMs-zero-to-hero：**这时候再看，你会发现你几乎能看懂所有内容了。
    - **Andrej Karpathy的"Let's build GPT: from scratch" YouTube：**这是LLMs-zero-to-hero的作者做的更详细的视频，他会一行一行地带你用PyTorch写出一个GPT。**这是你实现目标的终极指南！**
    - **开源项目：**在GitHub上搜索 "nanoGPT" 也是Karpathy写的 或 "minGPT"，这些都是简化版的GPT实现，代码量不大，非常适合阅读和学习。
-

### 第三部分：具体执行计划（时间表）

现在，我们把上面的路线图，填入你提供的时间块中。

**当前暑假（剩余14天 + 军训7天）**

**目标：点燃兴趣，建立全局观，启动编程基础。**

- **Day 1 今天 !:**
  - **任务：**花1.5小时，观看 **LLMs-zero-to-hero**。不用纠结细节，感受一下大模型的魅力，知道你的终点在哪里。
  - **任务：**花1.5小时，浏览 **廖雪峰的Python教程** 目录，安装好Python环境和Jupyter Notebook，跑通第一个"Hello, World!"程序。
- **Day 2 - Day 14 每天3小时:**
  - **1.5小时 编程：**系统学习 **廖雪峰的Python教程**。每天学习2-3个小节，并完成所有练习。目标是在14天内学完Python基础部分。
  - **1.5小时 数学：**观看 **【3Blue1Brown】《线性代数的本质》**。这个系列只有14集，每天看1集，反复琢磨，直到你真正理解“向量”、“线性变换”、“矩阵”等概念。这是你未来理解AI的基石。
- **军训7天 每天1小时:**
  - **任务：**休息很重要！这1小时用来复习和巩固。可以拿出手机，在B站上重看一遍《线性代数的本质》里你觉得最酷的一集，或者在脑子里回顾一下Python的语法。**不要学新知识，保持感觉即可。**

**高一全年（上课后每周1小时）**

**目标：稳扎稳打，完成阶段一。**

- **每周1小时安排：**
  - **前20周：**继续啃 **《Python编程：从入门到实践》** 的后半部分项目。每周花1小时，跟着书做项目。遇到问题，周末再花点时间解决。目标是完成至少1-2个项目。
  - **后20周：**开始学习 **NumPy**。找一些NumPy的入门教程（B站、CSDN上有很多），每周学习1-2个核心功能（比如数组创建、切片、运算）。目标是能用NumPy进行基本的矩阵运算。
- **寒暑假：**
  - **寒假/暑假 每天3小时，共约60小时:**
    - **1.5小时 编程：**主攻 **NumPy和Pandas**。找一些实战数据集（比如泰坦尼克号生还者预测），用Pandas做数据清洗和分析，用Matplotlib做可视化。
    - **1.5小时 数学：**主攻 **【3Blue1Brown】《微积分的本质》**。同时，确保学校的数学成绩优秀。

**高二全年（上课后每周1小时）**

**目标：完成阶段二，进入阶段三。**

- **每周1小时安排：**
  - **全年：**开始 **吴恩达《机器学习》课程** 20h。每周看1-2个视频，做好笔记。这个课程信息密度很大，慢慢来，不求快，只求懂。作业可以暂时放一放，或者用Python简单实现。
- **寒假** 每天3小时，共约60小时：
  - **核心任务：**完成**吴恩达《机器学习》课程**！每天集中精力看课、做笔记、思考。用剩下的时间，对照 **黑马程序员** 的课程，把重要的算法用Python+NumPy实现一遍。
- **暑假** 每天3小时，共约60小时：
  - **核心任务：**完成**吴恩达《深度学习》专项课程**！同时，并行学习 **PyTorch**。学完一门课，就立刻用PyTorch去实现里面的网络。目标是暑假结束时，能用PyTorch独立完成一个MNIST手写数字识别项目。

李沐

## 高三寒假（最后的学习时间）

**目标：完成阶段四，实现你的“大模型梦”。**

- **寒假** 每天3小时，共约60小时：
  - **核心任务：**跟着**Andrej Karpathy**的教程，**从零构建GPT**！
  - **Week 1:** 反复观看"Let's build GPT"视频，理解每一行代码的含义。
  - **Week 2:** 自己动手，一行一行地把代码敲出来，调试，运行。
  - **Week 3 & 4:** 尝试用自己的小数据集训练模型，或者修改模型参数，观察结果的变化。阅读nanoGPT的源码，理解更工程化的实现。

---

## 最后的话：给你的几点忠告

1. **健康第一：** 不要为了学习熬夜，保证充足的睡眠和运动。大脑休息好了，学习效率才高。
2. **保持耐心，拥抱挫折：** 你会遇到无数看不懂的公式、调不通的bug。这太正常了！每一个AI专家都是这么过来的。把解决问题看作是升级打怪，每解决一个，你就变强了一点。
3. **不要追求完美：** 计划是用来指导的，不是用来束缚的。如果某个阶段你觉得跟不上，就放慢速度。如果某个部分你特别感兴趣，就多花点时间。灵活调整。
4. **多输出，多分享：** 尝试写博客、做笔记，把你学到的知识讲给别人听（哪怕是对着空气讲）。这是检验你是否真正掌握的最好方法。
5. **享受过程：** 你正在探索人类智慧最前沿的领域之一，这本身就是一件无比酷炫和快乐的事情。享受每一次“啊哈！”的顿悟瞬间吧！

同学，你的这份计划充满了雄心壮志，而我为你制定的这份计划，则是为了让这份雄心能够脚踏实地。这条路很长，但每一步都算数。三年后，当你真的能看着自己训练出来的小模型吐出像模像样的文字时，你会感谢今天这个下定决心、目光远大的自己。

**祝你旅途愉快，未来可期！** 如果在学习过程中有任何问题，随时可以再来找我。