

零基础的新手，如何入门机器学习？

一、机器学习入门浅谈

机器学习领域，又或者更大而化之的说人工智能方向，因为“阿尔法狗”等一系列的热门爆点话题，被推到了人前，受到越来越多人的关注。无论你是什么领域的工作者，都一定多多少少听说过它的名号。而与此同时，随着机器学习领域理论的成熟，越来越多的相关技术被应用于生活实践的方方面面，从事机器学习岗也成了很多计算机行业从业者的重点考虑。

在这个chat里，我想简单的和大家探讨一些新手选择机器学习方向时总会考虑的问题。

其实，认真的说来，我们很难在网上找到一个教程，完全贴合每个新手入门机器学习时的需求。原因很简单，每个人在涉足这个领域的时候，个人的知识积淀和教育背景不一样，同样，他们对于未来发展的心理预期不一样，这也导致了需求的多样化。

因此，在真正入行之前，我非常建议你们，能够对自己进行一个深度的剖析，发现自己的优缺点，从而能够查漏补缺，才能在这个行业走的更远，同时，也能从心底对自己的选择更有底气。

从我个人的角度出发，我选择机器学习领域，是理智选择后的结果。我从高中起就觉得程序员这个职业特别cool，但是阴差阳错本科选择了数学，研究生才转行成为半个“非专业码农”。我的教育背景决定了和大部分程序员相比，我在数学理论上有一定的优势，但同时，我的代码能力非常的弱（在本科时，我只简单接触过python、matlab和c++）。我们很容易在学习中发现，机器学习方向对代码的要求，要远远小于传统的IT行业，比如前端、后台，甚至一些更底层层面的业务需求。综上所述，我才最终决定选择了这个方向。我个人觉得，我的这个考虑，更适用于还在上学的学生，或者刚刚毕业的毕业生们。

力是对现在的工作和薪酬情况不满意，同时被现在越来越火的机器学习领域所吸引，渴望改变现状。他们的工作量是非常巨大的，理论和基本代码能力两手抓，同时还需要补很多计算机的专业课程，属于三类人中困难度最高的。

在下面一部分中，我会简单总结一下我手头使用过或者比较了解的机器学习资料，并对他们进行点评和分类，以供不同侧重点的大家可以选择。

二、机器学习资料

写在前面的话：无论你是我刚才提到的哪一部分入门者，**数学基础**都是一定一定必要的！不认真打好基础的人，永远也不会爬得太高。

线性代数、高数和概率论与数理统计，是三本非常非常非常必要的课，如果忘记了请随时温习，直接搜最经典的版本教材就可以。

视频资料

我个人非常喜欢的方式，现在的网络这么发达，给我们了一个机会可以接触到全球各地的名师。刚入门的时候，名师的讲解肯定比自己死啃书要好，所以推荐给所有时间不那么紧张的小伙伴。

1. Coursera上Ng开设的机器学习

这门课十分经典，基本算是机器学习的入门课了。吴恩达在课堂上使用的是Octave，但是用什么语言实现其实都是可以的，课程的大作业，以及吴恩达写好的框架都对摸清行业基本工程框架十分有帮助的，基本可以算是理论与工程都兼顾。

相信大家都知道，Ng现在离开百度自己出来创业，普及深度学习，又推出了更新的深度学习几门课程，现在国内的一些主流网课app都可以找到课程资源。还是延续了老吴的一贯风格，更侧重理论一些，对代码的要求没那么高，对代码小白十分友好。

2. 斯坦福公开课cs231n

这部分很多都更侧向理论部分，毕竟任何方向要想走得足够远，就一定会涉及到晦涩的理论部分，这个时候，一本书和一本笔记，就是非常必要的了。

1. 南大周志华的《机器学习》

非常非常经典的西瓜书，基本上搞机器学习的人手一本。个人感觉属于典型的科班教科书，如果不是正在上学的学生，不是特别推荐，会有更多更好的替代读物，虽然名气大，但是不适合入门读，比较适合已经有基本概念的人来看。

2. 《deep learning》中文叫做《深度学习》

这本书具有极高的价值性和权威性，非常推荐，借用网上的评价大概是“这本书的价值在于让你低成本（高效率）的补上最近几年的主要理论成果”，非常适合面对数百篇论文无从下手的科研人士入门。当然也有人反应它的核心章节讲主流学习模型CNN和RNN的时候逻辑有点混乱，但是现在CNN和RNN发展到现在，不同的方向早就有了差异巨大的变种，所以，指望一本书能够全部搞明白也是不现实的。

3. 《Hands On Machine Learning with Scikit Learn and Tensorflow》

非常可爱的蜥蜴书，是某段时间相关主题的top1，和我接下来介绍的一本《机器学习实战》各有优劣，大家可以酌情选择。

这本书有很多深度学习在工程应用上的小窍门，解决了很多工程开发中的现实问题，但是同时，它的代码量没有机器学习实战那么多，理论讲解体系也相对更完善。非常适合刚上完NG课程的新手来衔接使用。

4. Machine Learning in action 中文名为《机器学习实战》

也是非常经典的一本书了，小伙伴基本人手一本，我从他还没出版的适合就开始关注。主要特点就是简单粗暴，特别适合和我一样理论还可以但是代码能力基本为零的程序员小白。“这本书用尽量少的公式把机器学习的基本算法都过了一遍，而且还讲得很清楚，更为重要的是他将公式和代码结合了起来。因此，你的机器学习并没有那么的抽象了，你知道算法里的公式如何的转化为代码。”（这段摘自知乎）

如果你的代码基础比较好，可以简单过一遍，缺乏代码量，对python等语言不熟悉的小

可能看到中间，就有小伙伴开始疑惑，欸，不是说机器学习吗？怎么不知不觉就说到了深度学习？

其实，人工智能领域发展到现在，机器学习已经远远不能满足我们的日常需求了。甚至可以这样说，机器学习的知识只是这个领域的冰山一角，我们仅靠单薄的机器学习的知识，早就不能找到一份理想的工作。计算机这个行业，从来都是活到老学到老，人们眼中的高薪，也就决定了它残酷的竞争性和知识迅速的更替性。

机器学习入门以后，就需要你根据自己的需求来制定个性化的路线了。

比如大数据分析、图模型、深度学习、自动驾驶、机器人、自然语言处理、图像识别、语音识别等等，既有应用方向，也有理论方向。正如我开篇所提到的，早在不知不觉间，机器学习就已经深入到了我们生活的方方面面。

在推荐中，我提到的都是自己亲自看过（虽然有一些并没有看完）的资料，也有很多名气很大很权威的资料没有提及，比如《Elements of Statistical Learning》《机器学习基石》等等等。毕竟个人能力有限，也欢迎大家一起来讨论补充。同时，我也坚信，现在这个时代，好的资料太多，而人的精力有限，与其把时间和精力花在筛筛选选上，不如静下心来把几本经典的教材吃透。

最后，无论你是从哪里开始，选择好方向，坚持走下去，才是唯一走到最后的诀窍，时间才是检验奋斗的唯一标准。

与君共勉。

GitChat