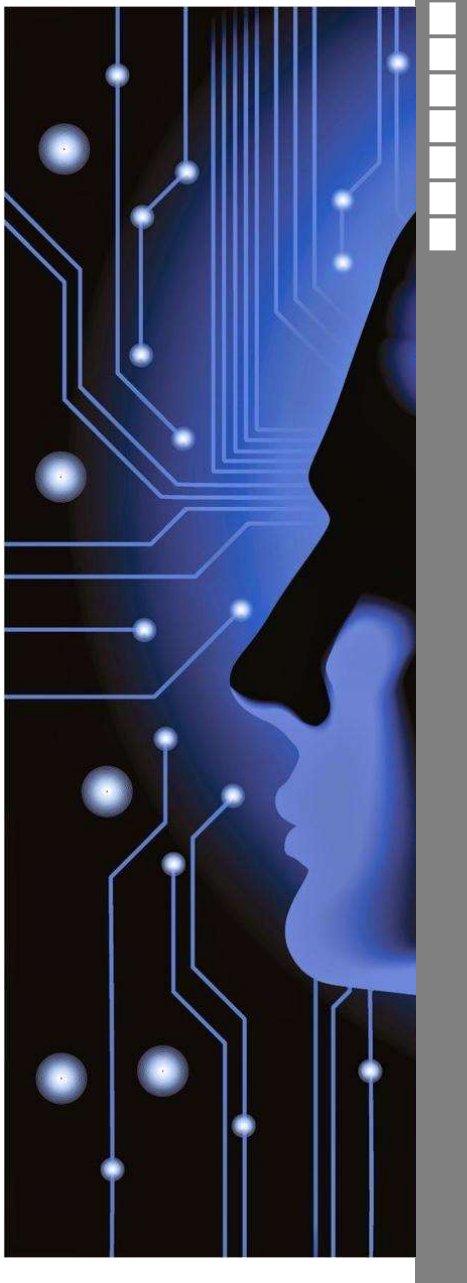
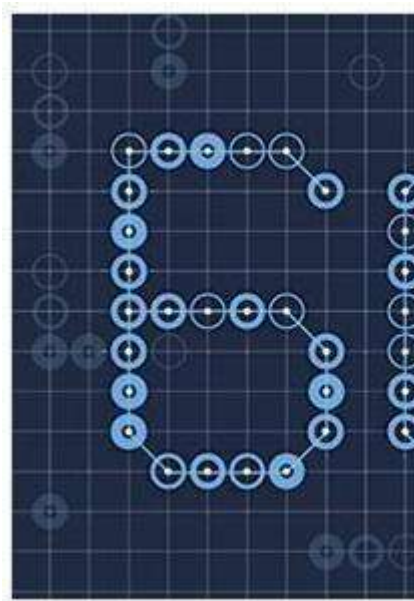


今天学习不努力，明天努力找工作



学习2017年1月，人类在智力领域数万年积攒下来的嚣张气焰遭到了若干次无情的打压。不提CES 2017上激动人心的自动驾驶产品（估计短短几年之后，你的驾驶证就可以扔掉了），也不细讲《最强大脑》节目里人类精英在图像识别环节被碾压（这曾经是人类可以嘲笑人工智能的典型领域），就说说围棋界的海啸吧。数十位人类围棋界顶尖高手，用人类围棋对弈的千年智慧结晶，历经艰苦卓绝的奋斗，依然被**Master**用快棋打了个一塌糊涂----60:0。



点击图片看原图

别人输棋以后都很茫然，唯独古力九段战后的话最为耐人寻味----“人类与人工智能共同探索围棋世界的大幕即将拉开”。

这话我听着，觉得韵脚极为熟悉。突然想起来王朔的《顽主》中描述过类似情节。

马青兴冲冲地走到了前面，对行人晃着拳头叫唤着：“谁他妈敢惹我？谁他妈敢惹我？”一个五大三粗，穿着工作服的汉子走近他，低声说：“我敢惹你。”马青愣了一下，打量了一下这个铁塔般的小伙子，四顾地说：“那他妈谁敢惹咱俩？”



玩笑归玩笑，不过人类精英终于学会谦逊，并对人工智能表达了合作的诚意，这也是好事儿。人工智能的崛起让很多人自以为认识到了一个现实：你今天有工作的时候不学习人工智能，明天人工智能会取代你的工作。

你同意吗？

其实这话错了，它不是现实。

人和机器不会发生竞争。人挣钱吃饭，机器有电就行。

现实是：你今天有工作的时候不学习人工智能，明天就有人利用人工智能取代你的工作。

人和人永远免不了竞争。这场军备竞赛里面，人工智能已经成为了个体可以掌握的核武器。别人懂了，你不懂，结果会很美妙。

当然，我说的是别人的结果。

人工智能领域中，目前最火的技术是机器学习(machine learning)，尤其是近年来伴随大数据产生而崛起的深度学习(deep learning)。不论是AlphaGo还是百度大脑，其实都是使用了（深度）机器学习，才变得如此厉害。

想不想学习一下“机器学习”呢？

途径要了解和掌握机器学习，路径有很多。你可以选课、报班，甚至自己看书。但是我更推荐的，是MOOC。

对于大多数人来说，MOOC资源更容易获得。只要你有一台能上网的电脑，就可以开始学。

因为这几年机器学习很火，因此相关MOOC很多。有好事者就开始进行梳理总结，把数十门课程分成初级、中级、高级，每一部分都有十余门课程。然后罗列出来，供读者按部就班来学习。

我的评价是，他们用非常认真负责的态度干了一件糟糕透顶的事儿。

为什么？

因为你一股脑罗列数十门课程，不会丝毫降低学习的门槛，反而迅速减少了“从入门到放弃”的时长。

没错，看完了你的清单，读者就放弃了。他的结论是“机器学习这事儿看来跟我这辈子没关系”。

本文尝试做一件与众不同的事儿----精选少数课程，组合一个易于入门和快速应用的路径。

要达到这个目标，就需要从许许多多的MOOC里选择“最好的”课程。所谓“最好的”课程，需要满足以下条件：授课教师自身是牛人；讲授用心；课程设计符合学习规律。

有了这个标准，许多课程立即被排除掉了。我们精选了数门课程，分别带领初学者完成入门、进阶、应用和补课等过程。组合起来，就是一个完整的机器学习了解和掌握路径。

众口难调，我们不可能找出所有人都一致认同的“最好”课程。如果你觉得我列出的课程或者组合方式还不够好，不要立即抱怨。因为这不是一场PK。欢迎你推荐心目中更好的课程给我们大家。

入门入门课程推荐吴恩达(Andrew Ng)教授Coursera平台上的《机器学习》。



吴恩达教授，就是《最强大脑》里面带着人工智能来砸场子的这位仁兄。他是斯坦福大学教授，Coursera的联合创始人之一。前任百度首席科学家，曾全面负责百度研究院。

如果你在MOOC论坛上发问，网友们给你推荐的第一门机器学习课程八成也是这门。原因是门槛低，天花板很高。

所谓门槛低，是指不要求你多么熟练掌握某种语言编程技术，也不需要你的数学知识达到做完若干册《吉米多维奇数学分析习题集》的水平。



所谓天花板很高，是指该课程全面介绍了机器学习中的核心知识，并且扩展到若干不同应用领域。这种课程设计方式，非常值得我国大学来学习。在保证覆盖面的同时，只在几个关键知识点着力反复练习，使得学生学起来轻松，却真的掌握到了精髓。

我觉得这才是真正的大师水平。

之前为这门课程专门写过一篇文章《机器学习，你也行》。这里就不再赘述了。欢迎感兴趣的同学查阅该文。

学完这门课，你就可以自豪地宣称自己了解机器学习了。

进阶获得了吴恩达教授课程证书之后，千万要趁热打铁，掌握现今最为热门的深度学习。



所谓深度学习，就是利用深度神经网络进行机器学习。如下图就是利用深度社交网络辨别一张图片究竟是车、人，或是动物。



此处推荐的课程是Geoffrey Hinton教授的《神经网络与机器学习》。

为什么推荐这门课？

首先，你会发现这门课使用的软件工具与上一门课程一样，都是易于安装、便于掌握的Octave。你不需要去了解编程环境的各种配置，拿过来直接用就行。框架代码都已经提供，你只需要在重点环节进行修改和补充。

其次，课程默认你学过吴恩达教授的课程作为前导。因此二者没有什么重叠，避免重复劳动。直接高屋建瓴，继续向纵深推进。

第三，Hinton教授是深度学习领域公认的权威。



这位大叔最大的本领就是对事业有自己的思考和见解，不人云亦云。当年人工智能领域遭遇寒冬，资助都没了，实验室都维持不下去了。许多有才华的人都跳到其他领域去了，唯独他老人家总是觉得不断改进与发展中应该有机会，就留了下来。

事实证明这种坚持带来的收获是显著的。他比别人提早深耕了数十年，如今的学者在深度学习领域无论写什么新论文几乎都绕不开引他的文献。

听这样泰山北斗级别的人物给你授课，真的是一种幸福。对别人来说一种非常高深莫测的算法，读起来要么如同天书，要么击节赞叹，可在他嘴里无非是“哦，那一年我试了一下这个，不成；后来突然开窍了，在这里改变了一点点，然后就.....”有些人在讨论区留言说他的笑话讲得太有趣了。这是真的。不过你学起来的时候对这个事儿期望不要太高----作为外国人，你的生活环境和文化底蕴不同，笑话就别指望全部能听懂了。

学完这门课，别人要是再想拿深度学习的概念来忽悠你，就很困难了。

这话你，真的听懂了吗？

学完这门课，如果你拿深度学习的概念去忽悠别人.....是不应该的。

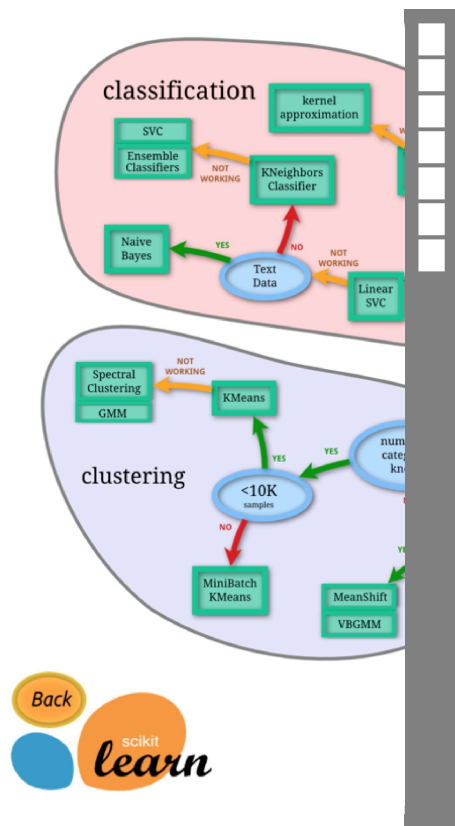
实践通过前面两门课，你理解了什么是机器学习和深度学习。但是你可能不愿意只停留在概念理解上，却很想做出些东西来实践。实践应用一方面可以检验你学习的效果，另一方面可以帮助你积累经验，同时还可以带来成就感。

不好意思，在应用实践环节这里，路径上有个分支。

第一种选择是Udacity上面的课程“Intro to Machine Learning”(ud120)。授课人是Sebastian Thrun，前几年他一直是Google自动驾驶技术的负责人。



这门课采用的函数库是**Scikit-learn**，构建于**Python**之上。课程设计得很吸引人，利用一些数据集，让你熟悉如何用别人构造好的功能（函数调用），加上你自己的理解（参数调整）来完成一些小型项目。



我对这门课程总体上很满意，但对免费学习用户，练习缺乏必要辅助，最后的大项目难度陡然提高。一般人难以适应，很容易中途放弃。如果你是付费用户，会有专人带领你一步步前行，并且提供必要的帮助，感觉就好多了。

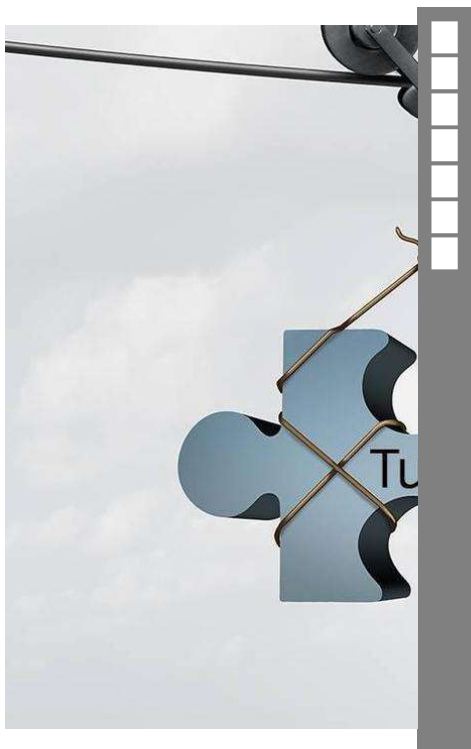
第二种是选择Coursera平台上华盛顿大学的机器学习课程专项系列(specialization)，课程名称为“构建智能应用”(Build Intelligent Applications)。



这门课程，充分体现出资本市场对于学术活动的干扰。

课程系列包含了6门课。但是只陆续推出了前4门。第5门一直在不停跳票。从2016年的9月份被反复推迟，后来干脆取消了。

为什么？因为主讲人两口子的公司Turi被苹果收购了。于是他们只好暂缓手头的课程制作，处理收购的各项事宜。



不过从已经推出的4门课来看，水平非常高，而且设计巧妙。

课程采用案例教学，就是每一个模块都依照具体问题给定数据集进行处理。第一门课是总览，让用户体会每一部分的重要内容，并且动手运行尝试，增加趣味和信心。

从第二门课开始，每门课分别介绍一个机器学习类型。每一周的课程由浅入深，逐渐加码。利用 **Jupyter Notebook** 保存练习的结果，将来可以直接用到实践中。

两位教授还非常大方地开放自己公司的产品免费给学生使用。他们的产品对于学术界也是相当于免费（每年需要手动免费更新一次证书），很慷慨。不过被收购以后，这个教育优惠政策的事儿就不好说了。

有的章节题目很有一些难度。作为学习者，你得做好远离舒适区，不停在学习区遨游的准备。

上述这两门课程选择一门听完，你应该可以具备自己动手处理一些小型数据集的能力了。对于学习、工作和科研应该会有很大的助益。

你很不满意，想要处理大规模数据？这个事儿需要两个条件支撑：你需要能得到大规模数据；你得有相应的硬件条件。

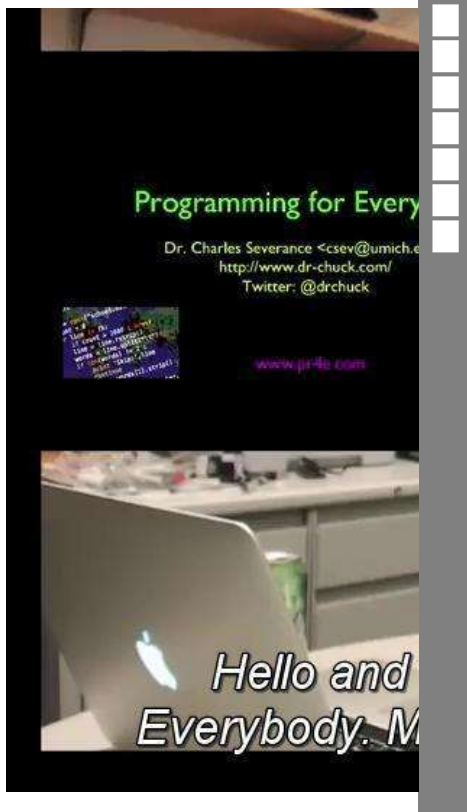
实话实说，目前绝大部分人不具备相应的条件。所以这个问题有空咱们再深入聊。

补课你可能对自己的知识基础没有什么信心----我不是数学或者计算机相关专业的，也能学会机器学习吗？

从刚才的课程介绍，你应该不难看出。现在机器学习的门槛已经很低了。但是为了达成目标，你还是需要掌握一些非常基础的知识，主要包括：数学（初级微积分、概率论与统计、线性代数）；编程语言（例如 **Python**）

如果你缺乏相应的知识，没关系，补就是了。缺哪里补哪里，反复补充，这样学习劲头和效率都有保障。

Python 语言的补课方式我建议采用密歇根大学的 **Coursera** 课程系列《人人都能学编程》。



这是目前我见到过的最适合数据科学初学者的Python语言入门方法。学了这门课程，你不仅掌握了一门语言，还可以掌握数据挖掘的基础知识。包括数据库操作、API和爬虫等。跟前面的实践类课程直接可以对接，严丝合缝。

至于数学知识的补充，我推荐的是Bill Gates一直赞不绝口的可汗学院。



可汗学院起家的本领就是帮助学生辅导数学功课。因此在MOOC领域，该平台的数学课程是积累最为丰富的，对学生来说也是最无痛的。



Math by subject

Early math

Arithmetic

Pre-algebra

Algebra

Geometry

Trigonometry

Precalculus

Statistics and probability

Calculus

Differential equations

Linear algebra

Math for fun

看到可汗学院这么多数学课程类别了吗？别犹豫，抓紧学吧。你不会重温大学一年级（或者高中三年级）初次涉猎微积分时的噩梦，因为Khan学院的讲授方式真的很生动、很有趣。讨论你还上过哪些讲解机器学习的好MOOC？它的独特之处在什么地方？能否替换本文列出的部分课程，形成更好的组合？欢迎留言分享，我们一起讨论。