

Yet Another Online Machine Learning Course

<http://freemind.pluskid.org/machine-learning/yet-another-online-machine-learning-course>

之前我们介绍过 [Coursera](#)，里面有内容丰富的各种课程，其中最早出现的就是 Andrew Ng 的 Machine Learning 课了。就在最近不久这个课程刚刚结束了它的第二轮（还是第三轮），如果错过了时间的同学可能要等到明年或者至少下个学期了。虽然任何时候都是可以观看课程视频的，但是如果不按照官方的进度来的话，就无法感受到课程作业的 **deadline** 大家一起学习所带来的动力和紧迫感了吧，而且作业似乎也不能在线提交了。

前几天还有个朋友跟我抱怨怎么那个课这么一下子就结束了，不过这档头有个好消息，就是网上又开了一门机器学习的课程。不过这次的课程不是挂在 Coursera 门下，但是也不是什么“山寨”货。课程链接在[这里](#)，虽然我也是今天才知道这个课程，但是毕竟是 Caltech 出品，应该是很不错的，课程网站上说的是“A real Caltech course, not a watered-down version”，也就是说这就是他们自己的机器学习课程，考虑到 Caltech 的名头（我想“看到我这篇 blog”、“没听说过 Caltech”和“没看过《The Big Bang Theory》”这三件事情同时发生的概率应该可以小到忽略不计了 :D），质量上应该是可以得到保证的。引用一下网站上对课程内容的介绍如下。

“*This is an introductory course on machine learning that covers the basic **theory, algorithms, and applications**. Machine learning enables computational systems to adaptively improve their performance with experience accumulated from the observed data. It has become one of the hottest fields of study today, with applications in engineering, science, finance, and commerce. The course balances theory and practice, and covers the mathematical as well as the heuristic aspects. The lectures follow each other in a story-like fashion.*”

形式上，也是看教学视频（和 Coursera 上一样是专门针对在线学习设计的视频，大部分时候是针对 slides 的展示，而不是简单的把普通的课堂录下来），也有课程作业并且是在线提交的形式，当然还有讨论组。不过视频是放在 Youtube 上的，所以 **blah-blah-blah**，然后讨论组似乎是另一个独立的论坛。这些细节上确实不如 Coursera 做得统一，不过我们就暂时不用去管这些细节了，何况 Coursera 也还有挺多毛病，比如他的那个在线播放器虽然能在课堂中间嵌入习题，但是用起来也不太好用，比如如果我想把视频往前 seek 半分钟重新看一下刚才说的什么的话，似乎就要重新缓冲一遍，卡在那里半天；而且有时候放着放着就只有声音了（不知道是不是 Linux 下特有的 bug），画面就定住了，因为画面经常都是静止的 slides 么，有时候过了很久才发现音频里讲的东西和画面已经完全对不上了..... =.=bb

然后，还有一点重要的是：7 月 10 日开始哦，现在已经开放注册了，

posted on [Free Mind](#) on July 7, 2012
generated with pandoc on December 3, 2015
category: Machine Learning

tags: Online Course, Resource

如果你还在为错过了上一次 Andrew Ng 的课程而后悔的话，千万不要再错过这一次机会了！:D 即使上过 Ng 的课的同学，应该也还是会从中得到收获的吧。对比起来的话，Ng 的课比较偏基础一点，而且介绍了许多 *intuitive idea* 和 *real-world practice*，例如课程最后讲到的在一个实际的 *Machine Learning Pipeline* 系统中如何去评测和规划应该将精力花在哪些部分，个人觉得是非常有用的，因为似乎我自己所在的实验室中的有些项目有时候就看起来在做一些比较“无用”的功——就是费了很大劲但是对整个系统的性能提升并没有太大的帮助。

而新开的这个课 (**WARNING**: 由于我也还没有上过这个课，所以下内容纯属梦话 :p) 因为说是 Caltech 自己用的 *Machine Learning* 课，所以从难度上来说应该会比 Ng 的那个课稍大一些，也会更深入一些，这一点从课程的 *schedule* 标题上也可以看到：这里还涉及到了 *Theory of Generalization* 和 *VC Dimension* 等一些内容，应该会比较有趣的。另外主页上的课程 *schedule* 实际上用颜色标出了每一课标题，三种颜色分别表示 *theory (mathematical)*、*technique (practical)* 和 *analysis (conceptual)*，可以作为内容分布的参考。

另外这个课程还有一本配套的课本，也是授课教授 Yaser S. Abu-Mostafa 自己写的书，叫做《*Learning From data*》，是今年刚出的新书（所以似乎目前既没有影印版也没有电子版.....:()，不过当我看到它只有 200 来页的时候就眼睛一亮了，大部头的书总是很难坚持读完啊！:D 顺便提醒一下，叫这个名字的书似乎有好几本，大家不要搞混淆了。

最后，趁这个档头我把机器学习相关的一些书搜集起来做了一个列表放在豆瓣上，主要放一些经典书籍（例如大家喜闻乐见的 *ESL*、*PRML* 等）和一些比较新的书（例如这里的这个课本，和 6 月份刚出的有两位 *Boosting* 之父联袂打造的新书《*Boosting: Foundations and Algorithms*》等），新书由于还没有经过时间验证是否是经典，所以收录进来的原因通常是作者很牛或者书的内容看起来很有意思等。一方面我会尽量避免滥竽充数地随便把什么书都加进来，但是另一方面我也不会拿什么来担保里面所有书的质量哦，况且有许多书我自己都还没有读过。

Update: 看了第一课的内容，觉得这个课应该会很有意思。lecture 长度大概在一小时左右，还有挺长的 *Q&A* 环节，至少第一次 lecture 里大家提的问题都是非常犀利的。课程的安排大致是按照这样的顺序来的：

1. What is Machine Learning?
2. Can we do it?
3. How to do it?
4. How to do it well?

此外，这里还有 John Langford 的推荐：

“Yaser points out some nicely videotaped machine learning lectures at Caltech. Yaser taught me machine learning, and I always found the lectures clear and interesting, so I expect many people can benefit from watching. Relative to Andrew

Ng *ML class*

”

hunch.net

另外 JL 在这里说的“Andrew Ng 的 ML 课程”又和 Coursera 上那个 ml-class 不太一样，前者比后者内容更多也更深一些，感兴趣的同学可以直接点上面的链接进去看一看。