

机器学习资料整理

写在前面： 做机器学习这么久了，以前也没有整理过这方面的资料，最近有同学问我要机器学习方面的资料，因此，整理了下我所知道的机器学习资料，希望能够帮助到想学习机器学习的同学。这些资料，有些我看过，有些是正在看的，有些是还没来得及看的。——by 贺成 2018.6.4 23:45

一、视频

1、 Andrew Ng 在 Coursera 上的《 machine learning 》，地址：<https://www.coursera.org/learn/machine-learning>。关于 Ng 无论再多的赞美也无法表达，就一句话吧：全世界人机器学习的入门导师（夸张修辞..）。课程浅显易懂，编程作业用 matlab/octave。

2、 Andrew Ng . CS229，这个是 ng 在斯坦福开的机器学习课，是 Coursera 上机器学习的加深版本。新学期的还没开课，没修过。地址：<http://cs229.stanford.edu/>

3、 Andrew ng. Coursera 上的《 Deep Learning Specialization 》，地址：<https://www.coursera.org/specializations/deep-learning>。ng 最新的课，这是个系列课程共 5 们，目前修到第四门，个人感觉最好没有之一。编程作业也与时俱进采用 python+jupyter notebook，再说一遍：讲的真的好。

4、 Geoffrey Hinton 《neural network for machine learning》，地址：<https://www.coursera.org/learn/neural-networks>。这门课当时只修到了第六周，就因为有些事情半途而废了，始终是个遗憾，准备有时间修完。老爷子讲的很好，但是稍微有点难（个人感觉），编程作业用的 matlab。

5、 林轩田 《Machine Learning Foundations》和《Machine Learning Techniques》。地址：<https://www.csie.ntu.edu.tw/~htlin/mooc/>。母语教学，只看了一点，不敢乱评价。

6、李宏毅《Machine Learning and having it deep and structured》和《machine learning》。地址：<http://speech.ee.ntu.edu.tw/~tlkagk/courses.html>。讲的真的是非常非常好。

7、徐亦达 《机器学习知识》，地址：
<http://i.youku.com/i/UMzIzNDgxNTg5Ng==?spm=a2hzp.8253869.0.0>。看了他的概率图部分，讲的也蛮好的，中文教学。

8、斯坦福 CS224d 《Deep Learning for Natural Language Processing》，地址：
<http://cs224d.stanford.edu/>。还没来得及看，不敢评价。

9、斯坦福 CS231n 《Convolutional Neural Networks for Visual Recognition》，地址：
<http://cs231n.stanford.edu/>。FeiFei Li 老师的课，和图像结合。还没来得及看，不敢评价。

10、莫凡的一些教学视频，适合入门的看。地址：
<https://morvanzhou.github.io/>。

二、书籍：

1、周志华 《机器学习》（西瓜书），浅显易懂，中文书最好的，我认为没有之一。。

2、李航 《统计学习方法》。偏公式推导些。

3、Ian goodfellow、bengio 等《deep learning》，地址：
<http://www.deeplearningbook.org/>。张志华老师团队翻译的中文版：
<https://github.com/exacity/deeplearningbook-chinese>。中文版已经出版，京东、当当、亚马逊都可以买到。

4、Peter Harrington 《机器学习实战》[Machine learning in action]。感谢蒋狗大二时花了 50 多大洋送我这本书。

5、范淼，李超 《Python 机器学习及实践：从零开始通往 Kaggle 竞赛之路》。个人感觉还不错。

6、Pang-Ning Tan 等著、范明等译《数据挖掘导论》[Introduction to Data Mining]，当年的启蒙书。。

7、袁梅宇 著 《数据挖掘与机器学习 WEKA 应用技术与实践（第二版）》，确切来说，个人感觉是翻译的，因为原版在：
<https://www.cs.waikato.ac.nz/ml/weka/book.html>。但是内容写的还是蛮不错的。

8、吴军 《数学之美》，大二时看到后再也无法忘怀，以致于买了一本送女票，后来自己又买了一本，再后来就是把身边的朋友洗脑都买了。。

9、Andrew ng 《machine learning yearning》，还未出版，电子稿见：<http://www.mlyearning.org/>（邮箱注册下就可以接受 ng 的手稿，目前还没写完）。这本书与其他的机器学习类书最大的不同就是这本书面向工程师，是 ng 多年的工业界经验的结晶。

10、Toby Segaran 《集体智慧编程》，很早就入手了，一直没时间看，不敢评价。

11、项亮 《推荐系统》，内容讲的很好，唯一的缺点就是这本书根本没有源码，特意和原作者交流过，这本书根本就没有代码。。不知道书上的代码是随手写的还是什么。但是概念介绍的还是不错的。

12、Christopher M Bishop 《Pattern Recognition and Machine Learning》（PRML），机器学习的经典之作，略难，目前只看了一点。

不是那些相关的书，但是个人感觉很好的：

1、吴军《浪潮之巅》，科普类最好没有之一。。强烈建议每个学计算机的都应该去读一读。记住：学科历史即学科本身。

三、竞赛

1、kaggle：<https://www.kaggle.com/>

2、DataCastle：<http://www.dcjingsai.com/common/cmptIndex.html>

3、阿里 天池比赛：<https://tianchi.aliyun.com/>

4、CCF 大数据：<http://www.datafountain.cn/#/competitions>

四、一些网站资料：

1、UFLDL：http://ufldl.stanford.edu/wiki/index.php/UFLDL_Tutorial

2、码农场：<http://www.hankcs.com/>

3、 Rickjin(靳志辉)《LDA 数学八卦》

4、 马晨《LDA 漫游》

5、 李宏毅 《一天搞懂深度学习》

五、常用工具

1、 scikit-learn, python 机器学习库

2、 weka, java 机器学习库

3、 TensorFlow , 谷歌深度学习库

4、 keras, 深度学习库

5、 pytorch , Facebook 深度学习库

6、 chainer , Intel 深度学习库

7、 caffe, 贾扬清, 深度学习库

8、 Mxnet, 李沐, 深度学习库

9、 paddlepaddle, 百度, 深度学习库

10、 jieba 分词, 中文分词

11、 FoolNLTK, 中文分词

12、 lightLDA, 微软

13、 lightGBM, xgboost, GBDT 竞赛三大神器。。。

14、 Numpy, 矩阵运算

15、 pandas, 数据处理

16、 networkX, 图, 社交网络分析工具

暂时想到这么多...