

我的机器学习入门清单及路线!

原创 桔了个仔 Datawhale 2022-03-01 21:00

Datawhale干货

作者: 桔了个仔, 南洋理工大学, Datawhale成员

知乎: <https://www.zhihu.com/people/huangzhe>

这是我个人的机器学习入门清单及路线，所以没有像很多收藏夹那样大而全，一来学不完，二来给自己压力。这是个人的路线。算是个人记录，也给大家参考，如有什么不足之处，欢迎指教。



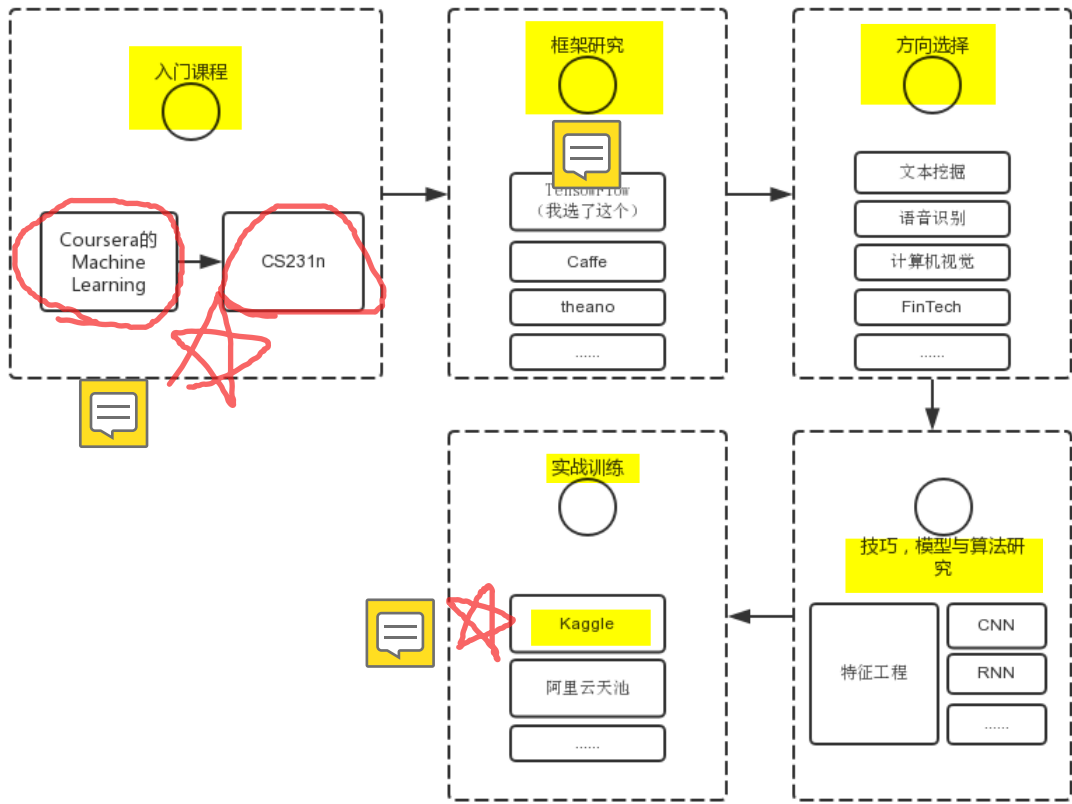
Datawhale
一个专注于AI领域的开源组织，汇聚了众多优秀学习者，愿景-for the learner，和学习者一起成长。
422篇原创内容

公众号

前置知识及技能:

- 1、**线性代数基础**，如果没的话，还是先学了这门课在研究吧，不然会哭的。
- 2、**学会python就行了**。R也可以有用。
- 3、**英语**。起码能基本的听和读吧，感觉中文的资料还不够多，很难避免要看很多英文资料。建议学习某些教程时看英文版的tutorial，YouTube可以开字幕。

做了个流程图，来展示下我的学习路线。



除了入门课程外，其他四项其实不完全是按照流程的（但总体上是），有时实战时需要学新模型。有时学了某些模型再选方向也未迟。但是入门课程，尤其是**Coursera那个**，一定要看完了才开始后面的学习。下面给出每项的学习地址：

入门课程



1. Machine Learning | Coursera 周志华机器学习

入门首选，推荐只认识“机器学习”四个字但还不知道它是什么的学习。这些年机器学习的大多数年轻人靠这个入门。具体提纲我就不列了，免得增加篇幅。建议是直接按顺序一课课学，不要着急。在学完这个课程前，不要学后面的。

链接: <https://www.coursera.org/learn/machine-learning>

2. CS231n: Convolutional Neural Networks for Visual Recognition



Stanford最受欢迎的课之一。做图像识别的，就算不想搞图像识别，也会学到很多有用的通用的东西。

链接: <http://cs231n.stanford.edu/>

框架研究



TensorFlow。推荐一个简单的中文入门教程。主讲人是一个在美国读大学的中国留学生，讲得很有趣。国内的朋友可以在bilibili看（一样的）

B站链接: <https://space.bilibili.com/16696495/#/channel/detail?cid=1588>

caffe

caffe是Facebook的开源框架。新的caffe2支持手机跑模型，可以说是非常先进了（当然手机性能嘛，就.....）比较好的中文教程没有找到，望有人留言推荐。

链接: <http://caffe.berkeleyvision.org/tutorial/>

theano

没怎么研究这个，但这个教程看起来不错

链接: <https://www.youtube.com/watch?v=OU8I1oJ9HhI>

技巧，算法与模型



特征工程：

特征工程到底是什么？ <https://www.zhihu.com/question/29316149/answer/110159647>

CNN:

卷积神经网络工作原理直观的解释？ <https://www.zhihu.com/question/39022858>

RNN:

循环神经网络RNN打开手册 <https://zhuanlan.zhihu.com/p/22930328>

你有哪些deep learning (rnn、cnn) 调参的经验？ <https://www.zhihu.com/question/41631631>

方向选择

机器学习领域，如何选择研究方向？ <https://www.zhihu.com/question/28689201>

深度学习目前主要有哪些研究方向？ <https://www.zhihu.com/question/23140232>

训练实战

阿里云天池：

天池大数据众智平台-数据科学家社区: https://tianchi.aliyun.com/?_lang=zh_CN

kaggle: 一个竞赛网站, 被谷歌收购后更火爆了。

地址: <https://www.kaggle.com/>

很多比赛是有奖金的, 能拿个奖基本大把公司主动给你工作机会。可以从Titanic这个比赛开始。

Titanic: Machine Learning from Disaster: <https://www.kaggle.com/c/titanic>

要了解kaggle, 我推荐一个不错的文章: <https://zhuanlan.zhihu.com/p/25686876>

计算平台:

有的时候我们需要带GPU的平台跑代码, 本地性能可能不是很多, 推荐几个(有免费有收费):

1. Kaggle. Kaggle可以用来当羊毛薅, 你可以传自己的dataset上去, 也可以用它提供的dataset.
2. Google Colab. 等于一个免费的notebook环境, 而且还能用GPU.
3. 百度AI Studio. 就是不支持tensorflow

篇幅虽然有点短, 但是学起来还是很久久的, 一起加油。

延伸阅读: 《Datawhale人工智能培养方案》



桔了个仔

南洋理工大学
Datawhale成员

Datawhale 和学习者一起成长

一个专注于AI的开源组织, 让学习不再孤独



长按扫码关注

整理不易, 点赞三连!

阅读原文

喜欢此内容的人还喜欢

极市CV技术开学干货! 从论文解读到算法实践、50+数据集资源、近3年顶会文章分类汇总
极市平台

机器学习需要掌握的九种工具!

好奇心Log

深度学习的发展方向: 深度强化学习!

数据派THU

