



Welcome to Beijing School of AI !

## 第二课：数据处理及SGD-机器学习



Siraj Raval  
Directeur, The School of AI  
AI Educator, Best-Selling Author, Youtuber

by Siraj Raval

SoAI项目Github地址



清华大学  
Tsinghua University

清华-青岛数据科学研究院

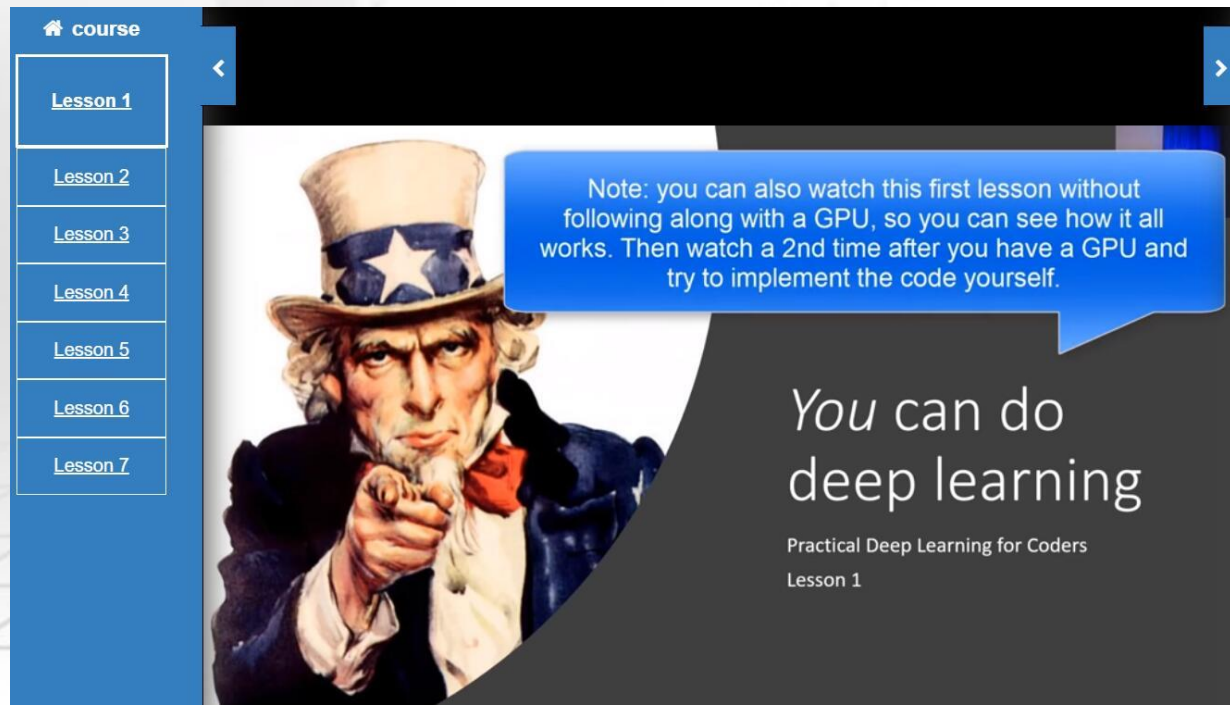
# 内容提要

- 一. 课程回顾-第一课
- 二. 内容简介-第二课
- 三. 集中讲解-问题解答
- 四. 机器学习专题
- 五. 总结预告

学 节	题 目	预习资源	正式解读
介绍	深度学习入门指南	请完成ppt里提到的作业, 长期	这节课长达2h, 内容非常多, 覆盖深度学习入门几乎全部知识点
第一 课	宠物图片分类,英文视频,中文 版笔记	Github代码	安装fastai环境并动手实现, 代码解析 Github本地, Colab, Kaggle
第二 课	特征工程及SGD,英文视频,中 文版笔记	Github代码	请提前预习左侧资源
第三 课	多标签分类,英文视频,中文版 笔记	Github代码	-
第四 课	NLP&推荐系统,英文视频,中文 版笔记	Github代码	-
第五 课	从反向传播到神经网络,英文视 频,中文版笔记	Github代码	-
第六 课	正则化卷积,英文视频,中文版 笔记	Github代码	-
第七 课	Resnets、GAN等,英文视频,中 文版笔记	Github代码	-

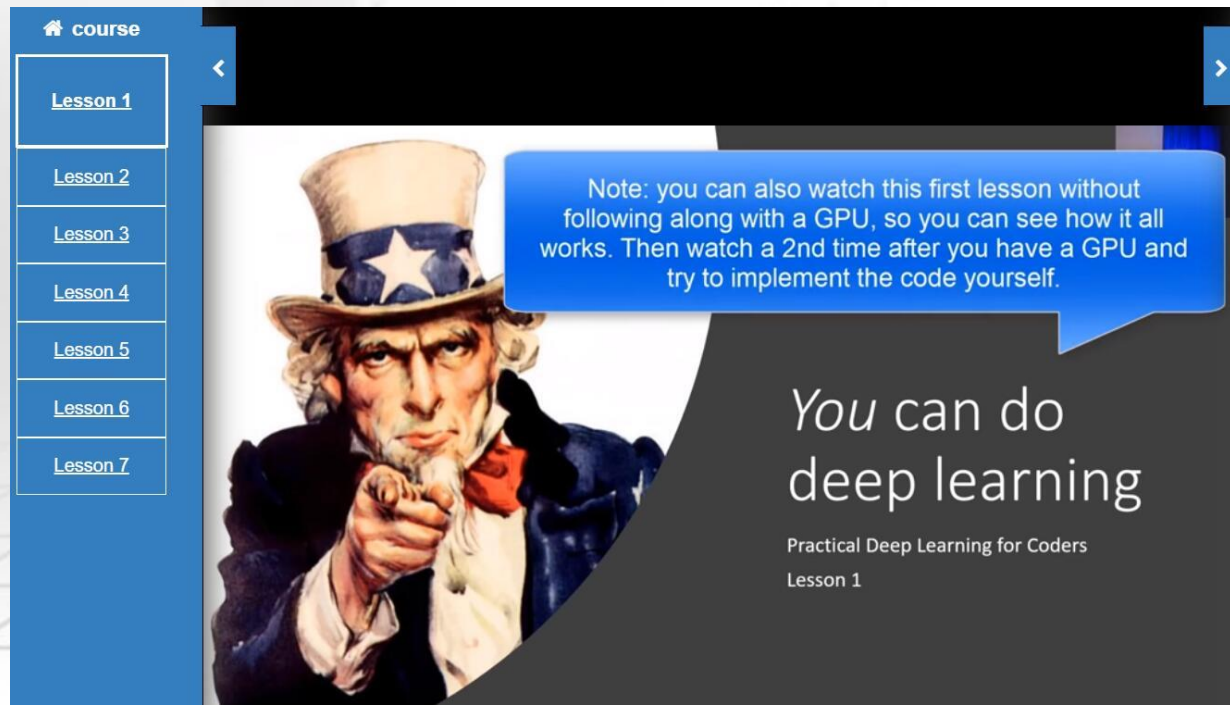
# 一 课程回顾

- 任务：图像分类
  - 动手实现一个图像分类 ([原版视频](#)、[国内源](#)、[Github地址](#))



## 二 内容简介

- 任务：数据集准备+手推SGD
  - 动手实现一个图像分类 ([原版视频](#)、[国内源](#)、[Github地址](#))



## 二 分组讨论

- 课程概要: <https://github.com/wqw547243068/school-of-ai-beijing/blob/master/README.md>
- 要点
  - 成员: 成员背景、编程功底、意图、组长
  - 温故: 上次课讲了什么, 印象最深的几点
  - 知新: 分享本次自习心得
  - 思考: 遇到哪些问题, 怎么解决
- 注意
  - 时间控制在10min以内
  - 每个子类目不超过3个问题
  - 尽量不要重复——先到先得
- 组长
  - 协调、监督组员进度
  - 结伴学习, 共同成长





## 三 集中讲解

- 知识点

- 机器学习流程
- GPU
- 图像处理
- 卷积神经网络
- . . .

Black box

- Interpretable ML
- Visualize gradients and activations

Needs too much data

- Transfer learning
- Share pre-trained nets

Needs ML PhD

- No longer true
- fastai & keras libs, MOOCs, etc

Only for vision

- No longer true
- SoTA for speech, structured data, time series...

Needs lots of GPUs

- Was never true
- ...except for some research projects

“Not really AI”

- Who cares?
- Do you really want to build a brain?



# 如何高效学习

## 互联网时代学习之道：

### • 多看：知识图谱

- 系统阅读：结构生长
- 碎片阅读：开枝散叶

### • 多动手：验证

- 消化理解，提升留存
- 开花结果，学以致用

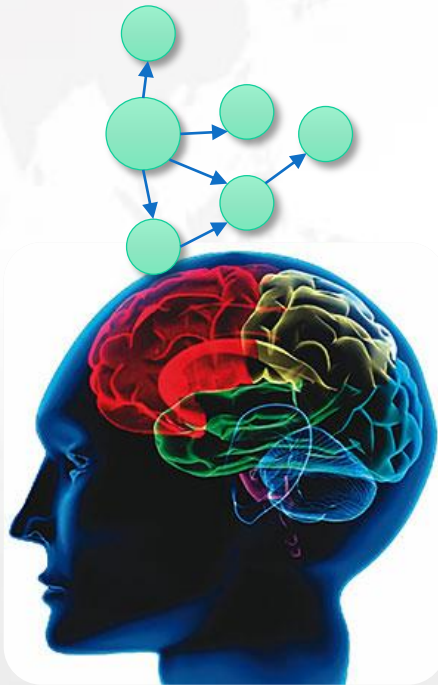
### • 多思考：关联推理

- 提炼关联，查缺补漏
- 不断完善知识图谱
- 沉淀：自己做笔记

### • 其他：好奇心+分享+上进心

## 大脑的学习之道：

### • 图谱结构+注意力+联想记忆+推理反思

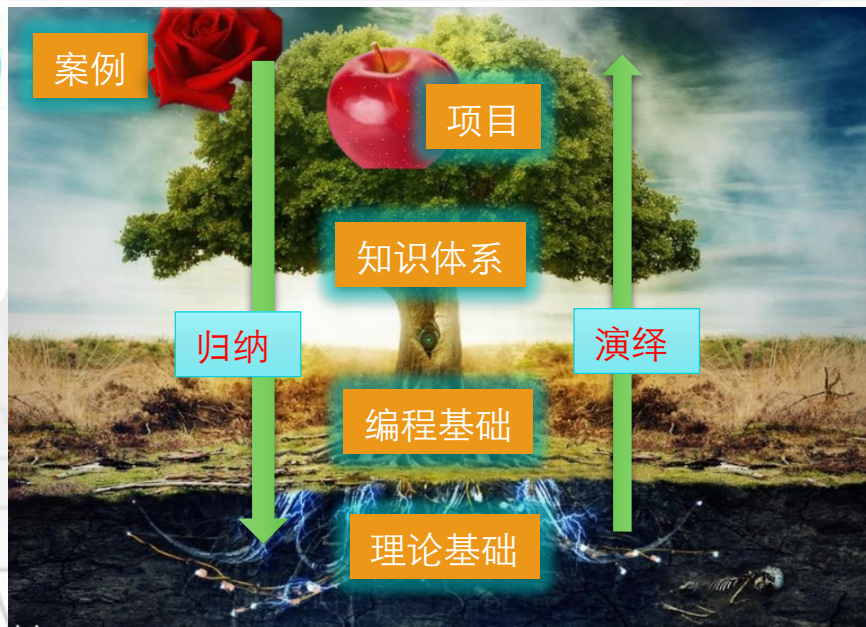


@爱可可-爱生活

读书重在结构生长，形成扎实的支撑；碎片阅读重在视野的纳新和扩展，开枝散叶；思考重在提炼和关联，勾画错综的经脉。学习就是如此，由外而内，无广不精，无博不深，但能坚持必有所成。网络阅读的最佳实践，不在“取”，在“舍”，知舍才能知关键，料不在多，有感悟一二足矣。

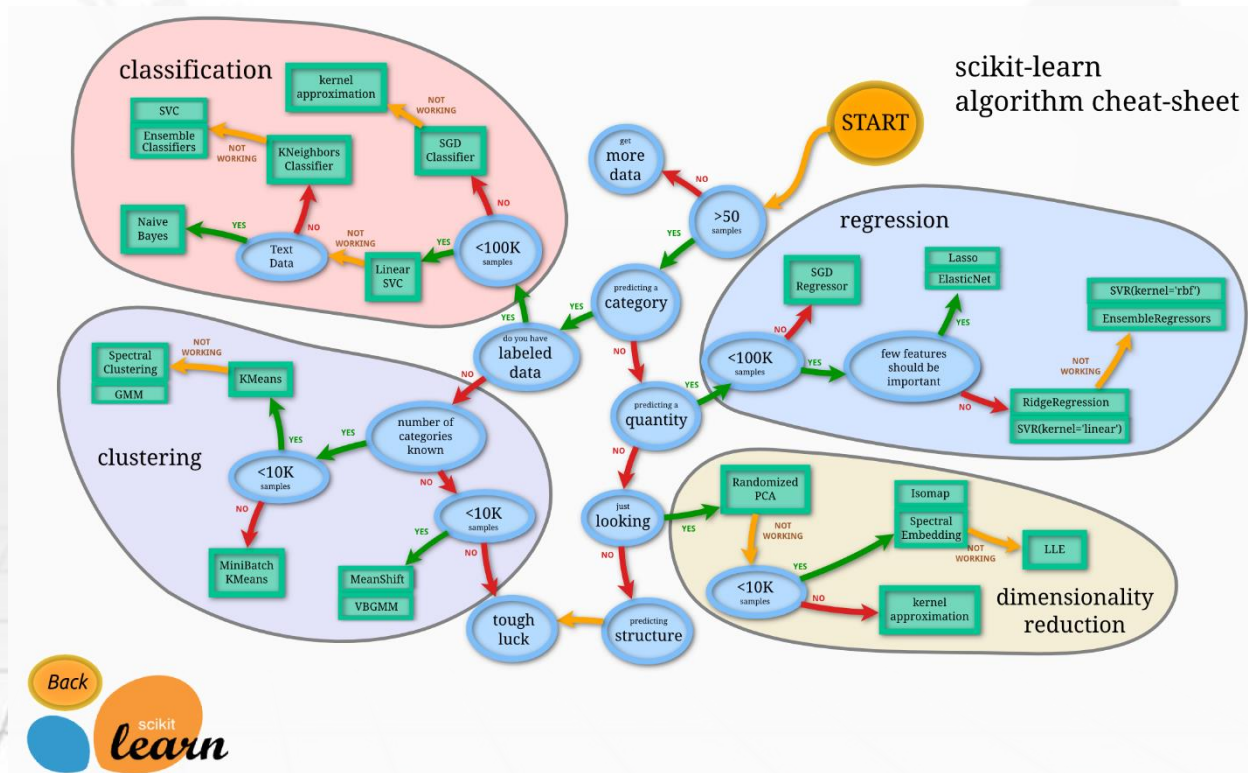
2015-6-16 10:37 来自 微博 weibo.com

1657 | 271 | 2407



# 四 机器学习专题

- 内容
  - 第一次课内容





# 五 预告

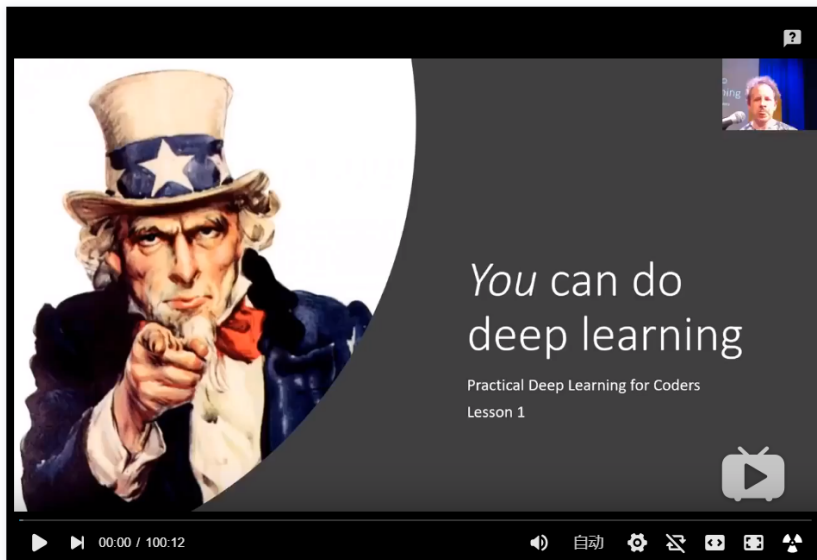
- 第三课 神经网络与深度学习
  - [哔哩哔哩地址](#)

[SoAI项目Github地址](#)

面向程序员的深度学习实战课程(2019)(PyTorch/fastai)

科技 > 演讲 公开课 2019-01-26 10:39:17

1.1万播放 · 1弹幕



fly51fly 发消息

没错, 是我 - 新浪微博 @爱可可-爱生活 http://...

已关注 4.5万

弹幕列表

展开

视频选集

1/7

P1 Image classification

P2 Data cleaning and production; SGD from scratch

P3 Data blocks; Multi-label classification; Segment...

P4 NLP; Tabular data; Collaborative filtering; Emb...

P5 Back propagation; Accelerated SGD; Neural n...

P6 Regularization; Convolutions; Data ethics

P7 Resnets from scratch; U-net; Generative (adve...

相关推荐



【公开课】深度学习与计算机视觉【19集】

庄七

8.7万播放 · 332弹幕

深度学习项目实战 (下)

# 翻转课堂

## • 约法三章

- 随机分组：全部学员5-6人一组，随机分配，尽量均衡，组团学习
- 组内互助：选组长，督促学员学习，相互帮助
- 组间竞争：每次课会对表现优秀的组加分，动态排名
- 奖惩分明：最后一名自觉给第一名买奖品（零食、红包等）

## • 准备工作

- 分组：报数
- 去中心化：围着讲台散开
- 积分榜
- QQ学习群（左图）：**The School of AI**
  - 资料共享、作业发布
- **QQ讨论组**（右图）
  - 如果需要共享自己的屏幕，需要使用讨论组
- 添加小助手alpha微信：**xiniuedu5**（右上角二维码）



群名称: SoAI-北京  
群 号: 1019542361





# THANKS

王奇文-wqw547243068@163.com