



链滴

Andrej Karpathy | 详解神经网络和反向传播 (基于 micrograd)

作者: [matpool](#)

原文链接: <https://ld246.com/article/1663318538002>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

只要你懂 Python，大概记得高中学过的求导知识，看完这个视频你还不理解反向传播和神经网络核要点的话，那我就吃鞋:D

Andrej Karpathy，前特斯拉 AI 高级总监、曾设计并担任斯坦福深度学习课程 CS231n 讲师、OpenAI 创始成员和研究科学家。在 7 月离职特斯拉后，Andrej 在家录制了一个详解反向传播的课程，自信示“这是 8 年来领域内对神经网络和反向传播的最佳讲解”，并在推特打赌“看不懂就吃鞋”。



Andrej Karpathy @karpathy · Aug 17

!!!! Ok I recorded a (new!) 2h25m lecture on "The spelled-out intro to neural networks and backpropagation: building micrograd"

youtube.com/watch?v=VMj-3S...

This is the culmination of about 8 years of obsessing about the best way to explain neural nets and backprop.



youtube.com

The spelled-out intro to neural networks and back...

This is the most step-by-step spelled-out explanation of backpropagation and training of ...

228

1,419

8,368



Andrej Karpathy @karpathy · Aug 17

If you know Python, have a vague recollection of taking some derivatives in your high school, watch this video and not understand backpropagation and the core of neural nets by the end then I will eat a shoe :D

32

52

1,368

知乎 @请叫我Tony老师 @稀土掘金技术社区

虽然很想看 Andrej 直播吃鞋，但实话实说，作为小白的我，好像真的看懂了。

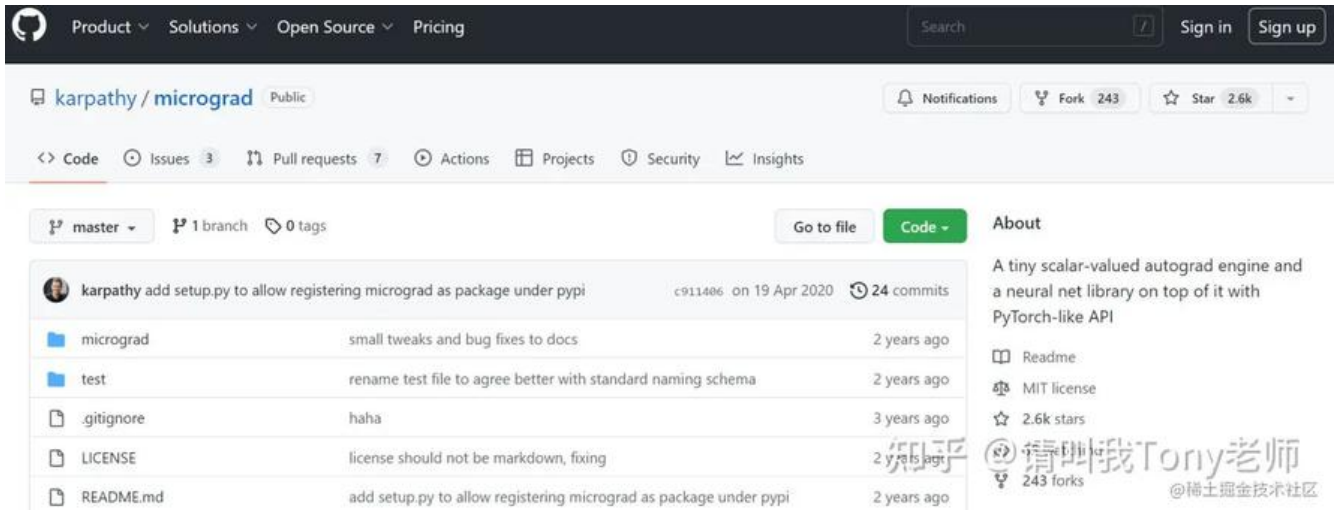
课程视频

[【中英字幕】Andrej Karpathy | 详解神经网络和反向传播 \(基于micrograd\)](#)

字幕由[矩池云](#)翻译制作

课程介绍

课程时长 2 小时 25 分钟，基于 micrograd 详细介绍并演示了神经网络的搭建和反向传播的过程。Micrograd 是 Andrej 于 2020 年编写并开源的微型 autograd (自动梯度) 引擎。其中用 100 行代码实现了针对动态构建的 DAG 的反向传播算法，并用 50 行代码实现了类 PyTorch API 的库。目前，micrograd 项目的 GitHub Star 量已达 2.6k。



基于 micrograd, Andrej 介绍了只有一个输入的简单函数的求导 ($y = 3x^2 - 4x + 5$ 这个函数中, 求对 x 的导数) 和具有多个输入的简单函数的求导 (已知3个输入 a 、 b 、 c , 并且 $d = a * b + c$, 求 d 别对 a 、 b 、 c 的导数)、启动 micrograd 的核心 Value 对象及其可视化的方法, 并举了两个手动反传播的例子。接下来 Andrej 系统地介绍了:

- 为单个运算实现反向传播;
- 为整个表达式图实现反向函数;
- 修复反向传播中一个节点多次使用出现的 bug;
- 使用更多运算替换 tanh 激活函数;
- 在 PyTorch 中实现上述运算并与 micrograd 进行比较;
- 用 micrograd 构建一个神经网络库 (MLP) ;
- 创建一个微型数据集, 编写损失函数;
- 收集神经网络的所有参数;
- 手动进行梯度下降优化, 训练网络。

最后 Andrej 总结了反向传播与现代神经网络的关系。此外, Andrej 还探究了 PyTorch 中 tanh 激活函数的反向传播机制

Chapters:

00:00:00 intro

00:00:25 micrograd overview

00:08:08 derivative of a simple function with one input

00:14:12 derivative of a function with multiple inputs

00:19:09 starting the core Value object of micrograd and its visualization

00:32:10 manual backpropagation example #1: simple expression

00:51:10 preview of a single optimization step

00:52:52 manual backpropagation example #2: a neuron

01:09:02 implementing the backward function for each operation

01:17:32 implementing the backward function for a whole expression graph

01:22:28 fixing a backprop bug when one node is used multiple times

01:27:05 breaking up a tanh, exercising with more operations

01:39:31 doing the same thing but in PyTorch: comparison

01:43:55 building out a neural net library (multi-layer perceptron) in micrograd

01:51:04 creating a tiny dataset, writing the loss function

01:57:56 collecting all of the parameters of the neural net

02:01:12 doing gradient descent optimization manually, training the network

02:14:03 summary of what we learned, how to go towards modern neural nets

02:16:46 walkthrough of the full code of micrograd on github

02:21:10 real stuff: diving into PyTorch, finding their backward pass for tanh

02:24:39 conclusion

02:25:20 outtakes :)

知乎 @请叫我潘金莲老师

课程主讲

Andrej Karpathy 出生于斯洛伐克，15 岁时随家人移民加拿大。在一次采访中，Andrej 表示，“在斯洛伐克的时候并不开心，一直想要离开，当父母表示要移民加拿大的时候，我立马同意了，但家里其人很犹豫，我就去说服每个人” [3]。

来到加拿大的 Andrej 还不会说英语，但数学成绩非常优秀，完成高中学业后进入多伦多大学计算机科学与物理系并辅修数学，在那里他遇到了 Geoffrey Hinton 教授，也是在 Geoff 的课上，他接触到深度学习算法。

2009 - 2011 年，Andrej 硕士就读于加拿大不列颠哥伦比亚大学，其导师为计算机科学系教授 Michie van de Panne，主要研究物理模拟中用于敏捷机器人的机器学习。

CS231n 初露锋芒

2011 年，Andrej 进入斯坦福大学攻读博士学位，师从李飞飞主攻深度学习和图像识别。在斯坦福，和导师李飞飞一起研究图像识别与深度学习算法，主导设计了斯坦福第一门深度学习课程《CS231n 用于视觉识别的卷积神经网络》，并和李飞飞共同担任主讲人。很快，这门课成为斯坦福的网红课程注册学生从开课时的150人增加到两年后的750人。

在读博期间，Andrej 发表的多篇论文都是 CV 领域的高引文章，还在 Google、DeepMind 两家公司实习。

他会是全世界最顶级的AI领袖

2016年 Andrej 成为 OpenAI 创始成员，OpenAI 是由马斯克等人创立的非盈利人工智能研究组织，是在这里，他的天赋被马斯克看中，马斯克不惜得罪 OpenAI 其他成员把 Andrej 从 OpenAI 带到了特斯拉。五年里，Andrej 一手促成了 Autopilot 的开发。随着特斯拉从最初的自动驾驶慢慢扩展到广泛的人工智能领域，他也被提为特斯拉的AI高级总监，直接向马斯克汇报工作。

据说马斯克曾说过，很多人把 Andrej 当做优秀的 AI 视觉科学家，但我知道，他会是全世界最顶级的 AI 领袖。

而这位 AI 领袖现在才36岁...