



TAIJI LABORATORY
FOR GRAVITATIONAL WAVE UNIVERSE



ICTP-AP
International Centre
for Theoretical Physics Asia-Pacific
国际理论物理中心-亚太地区



中国科学院大学
University of Chinese Academy of Sciences

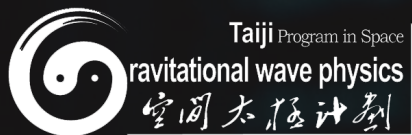
引力波数据探索：编程与分析实战训练营

第 0 部分 课程介绍会

主讲老师：王赫

ICTP-AP, UCAS

2023/11/08







目录

- 办课初衷与学员构成
- 讲师介绍
- 与本课程相关的知识架构
 - 引力波数据分析
 - 课程大纲
 - 本课程是什么，不是什么
- 本课程的学习方法与教学团队
- 本课程的考核规则和项目作业
- 通向自我实现之路
 - 如何自学
 - 如何提问
- 提问环节





办课初衷



引力波数据探索

编程与分析实战训练营

报名啦!

本课程面向数据分析与算法开发感兴趣的本科生或研究生，尤其推荐给引力波数据处理研究领域的热衷者，旨在提升学员在Python编程、运维知识及数据驱动模型建立的基础技能，同时深度探讨人工智能在引力波数据处理中的应用。

通过实例与项目使学员能够掌握数据驱动建模解决实际问题，为未来的空间引力波探测项目和相关研究领域培养高质量人才。

报名方式

报名时间：2023年10月27日至2023年11月5日
线上申请链接：<https://ictp-ap.org/form/54#/>

开班时间


拟于2023年11月8日开班，预计于2024年1月结束

授课方式

- 线上线下结合：线上--腾讯会议/线下--国科大中关村校区
- 每次授课时间约2-3小时，实操占80%
- 初步计划为每周三或周五19:00-22:00
- 提供录播回放，学员能够随时回顾和加深理解

课程大纲

第一部分：编程开发环境与工作流
第二部分：基于PYTHON的数据分析基础
第三部分：机器学习基础
第四部分：深度学习基础



扫二维码了解更多详情
或至：<https://ictp-ap.org/event/122#/> 查看

主办单位

中国科学院大学 国际理论物理中心（亚太地区）引力波宇宙太极实验室

Name	Last Modified
vt_mbhb217	21 days ago
vt_mbhb218	21 days ago
vt_mbhb219	19 days ago
vt_mbhb220	19 days ago
vt_mbhb221	19 days ago
vt_mbhb222	19 days ago
vt_mbhb223	19 days ago
vt_mbhb224	18 days ago
vt_mbhb225	19 days ago
vt_mbhb226	19 days ago
vt_mbhb227	19 days ago
vt_mbhb228	18 days ago
vt_mbhb229	17 days ago
vt_mbhb230	16 days ago
vt_mbhb231	17 days ago
vt_mbhb232	17 days ago
vt_mbhb233	17 days ago
vt_mbhb234	17 days ago
vt_mbhb235	17 days ago
vt_mbhb236	17 days ago



```
91 parser.add_argument('--use_gpu', default=False, help='If true, use gpu /
92 args = parser.parse_args()
93
94
95 global params; global bounds; global fixed_vals
96
97 # Define default location of the parameters files
98 params = os.path.join(os.getcwd(), 'params_files', 'params.json')
99 bounds = os.path.join(os.getcwd(), 'params_files', 'bounds.json')
100 fixed_vals = os.path.join(os.getcwd(), 'params_files', 'fixed_vals.json')
```

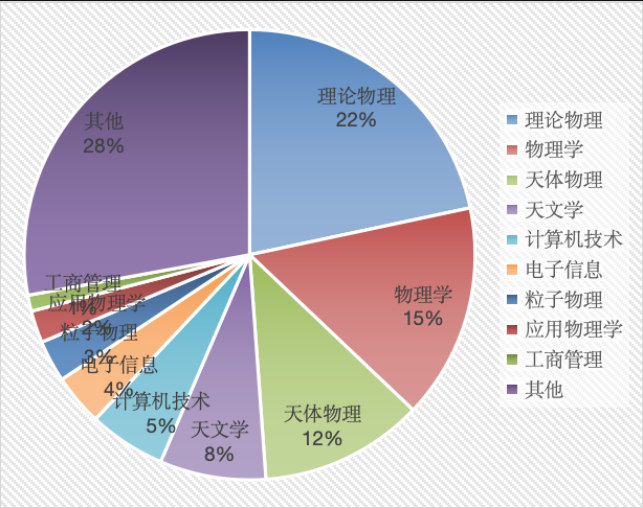
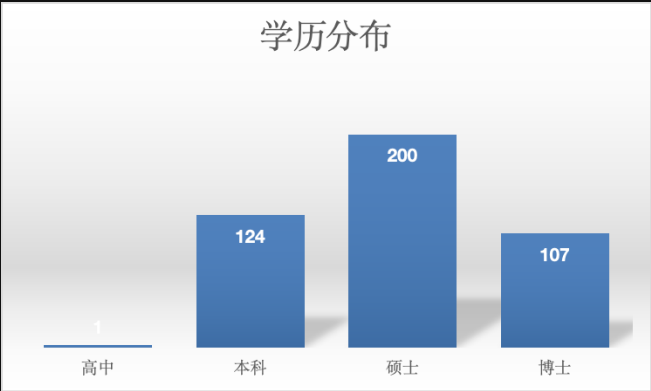
- 发现物理系研究生新生普遍缺乏基本的计算机常识，尤其是集群维护、Python编程、数据分析流程等等。
- 因此，期望能快速提升国科大新生的基本编程素养。



学员构成

- 有效申请的学员人数共 432 人，来自各行各业。其中一半以上来自理论物理、物理学、天体物理和天文学等专业方向。

- 在学员填写的“个人研究的主要障碍”和“对本课程的期望”的词云分析中，“引力波”、“机器学习”、“数据处理”、“深度学习”、“能力”、“编程”等是最常见的关键词。



- 在“其他”专业类别中，包含心理学、生物信息与医药、工商管理、材料与化工、地震与地球物理、金融、控制工程等等。





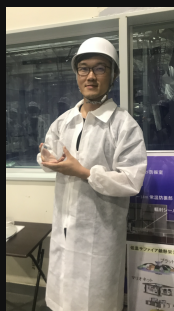
讲师介绍

王赫 He Wang,
LIGO-VTRGO-KAGRA 科学合作组成员。2020 年博士毕业于北京师范大学物理系理论物理专业，2020-2022 年于中国科学院理论物理研究所从事博士后研究，2021-2022 年赴鹏城国家实验室担任访问学者，2022 年至今为国科大国际理论物理中心（亚太地区）特别研究助理，2022 年获国家天文科学数据中心青年数据科学家称号。

GitHub: github.com/iphysresearch

Blog: IPhysResearch.github.io

- Aug 2023
 - 引力波暑期学校《引力波数据处理与机器学习》：中山大学天琴中心
- Jan 2022 - May 2022
 - 《数据可视化初了解》：知乎盐选
- Apr 2021 – Jul 2021
 - 人工智能从业者认证课程：腾讯云·课程开发
- Dec 2020 – Jul 2021
 - 《深度学习之 PyTorch 实战》：万门大学
- Jun 2020 – Nov 2020
 - 《Python 数据可视化与实战》：万门大学
- Apr 2020 – Oct 2020
 - 《大数据预处理》国家开放大学·大数据挖掘与技术系列课程
- Feb 2020 – Nov 2020
 - 《Python 程序设计》国家开放大学·大数据挖掘与技术系列课程
- Aug 2019 – Aug 2019
 - 《Python 数据挖掘工具》中原银行·校招生培训系列课程



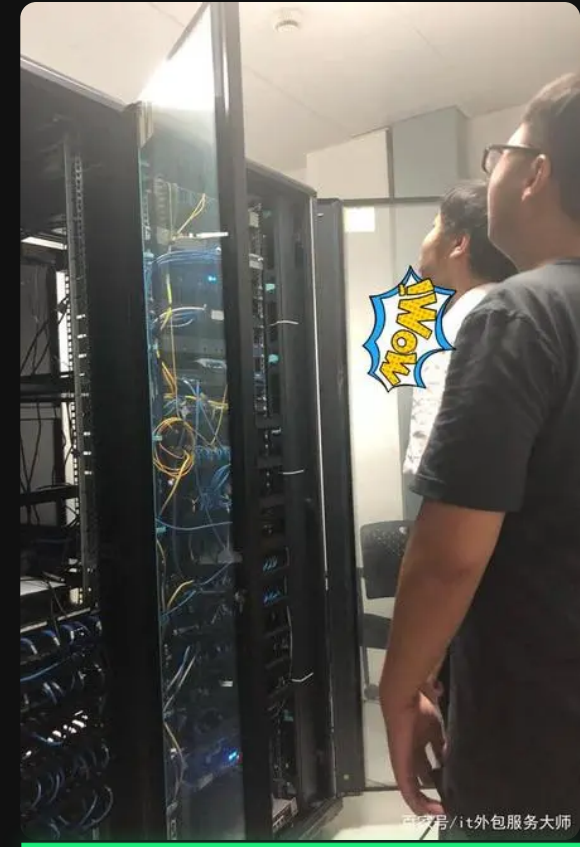
引力波数据分析?



程序猿?

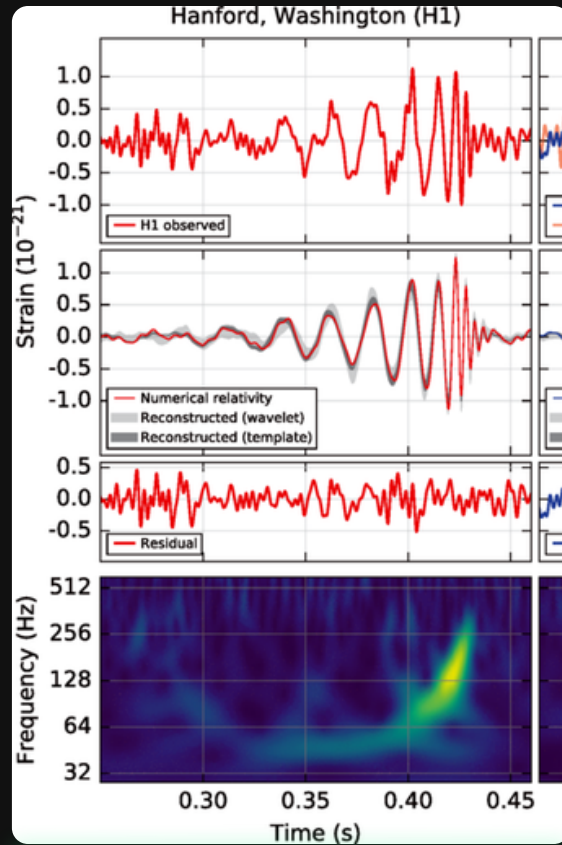


数据分析师?



运维工程师?

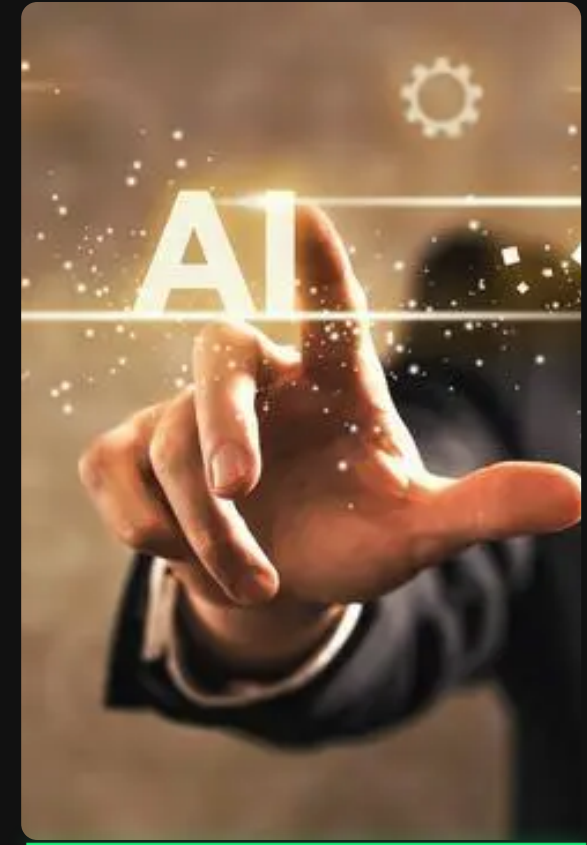
引力波数据分析!



引力波物理科学家



数据科学家



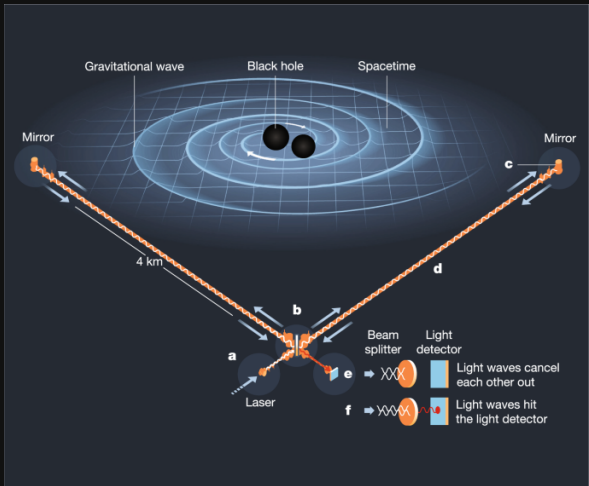
人工智能
算法工程师



引力波数据分析：知识框架

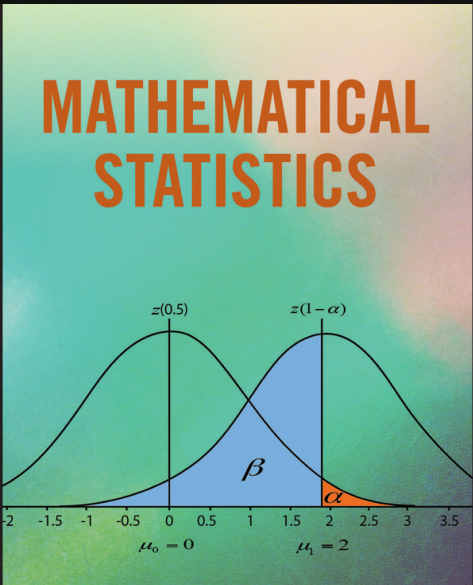
- 理论基础：
 - 引力波物理
 - 数字信号处理
 - 数理统计
- 编程基础：
 -
 -
 -
- 硬件基础：
 -

引力波物理与引力波天文学



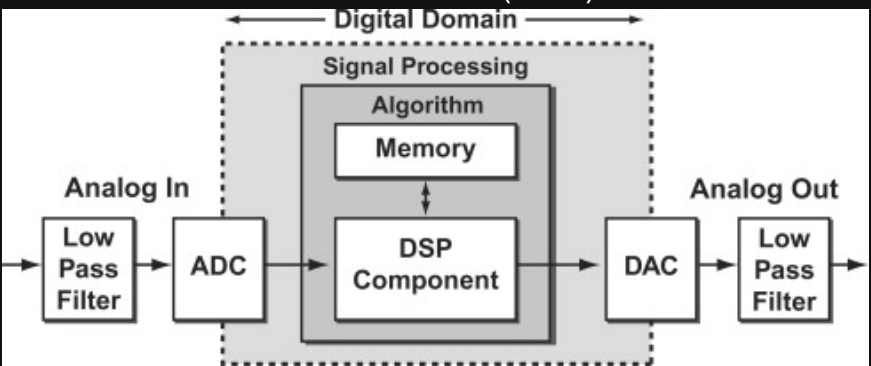
Miller, M.C., Yunes, N. The new frontier of gravitational waves. *Nature* 568, 469–476 (2019).

数理统计



Dieter Rasch, Dieter Schott. Mathematical Statistics, (2018)

数字信号处理 (DSP)

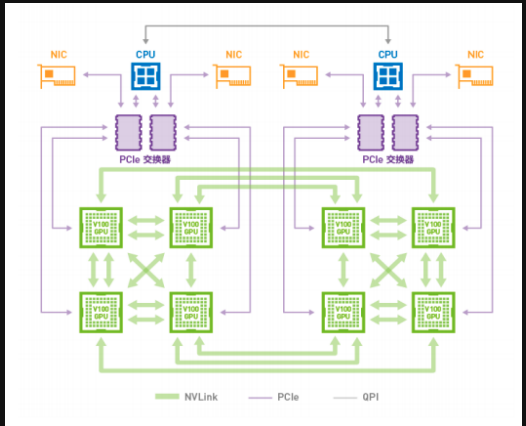
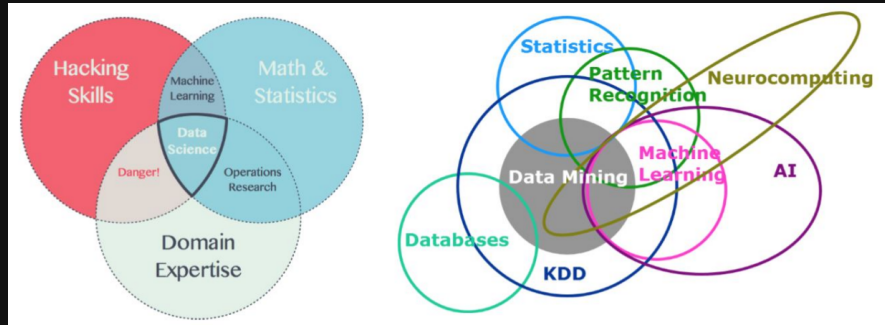


R.C. Cofer, Benjamin F. Harding, in *Rapid System Prototyping with FPGAs*, 2006



引力波数据分析：知识框架

- 理论基础：
 - 引力波物理 (pycbc, lalsuite, lisacode, bilby, ...)
 - 数字信号处理 (scipy, stat, ...)
 - 数理统计 (bilby, emcee, ptemcee, ptmcmc, ...)
- 编程基础：
 - Python (numpy, pandas; matplotlib; ...),
 - AI (scikit-learn, XGBoost, PyTorch, TensorFlow, JAX, ...)
 - Linux (docker, github, bash, vim, emacs ...)
- 硬件基础：
 - 主板、内存, GPUs, 显存 ...
 - AI时代程序员都应该了解的GPU基础知识





本实战训练营的课程大纲

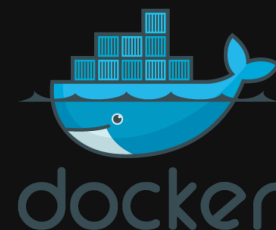
- 第零部分：打鸡血！
- 第一部分：编程开发环境与 workflow
 1. Linux 中常见指令与 shell 文本编辑
 2. Git 版本控制 (GitHub / GitLab)
 3. SSH 远程服务器 (Shell / VSCode)
 4. 基于 Docker 的容器化技术与开发环境
 5. 实战项目：Python / Jupyter 开发环境搭建
 6. 实战项目：LALsuite / LISAcodes 的源码编译
- 第二部分：基于 Python 的数据分析基础
 1. Python 程序语言入门
 2. 基于 Numpy / Pandas / Scipy 的算法基础
 3. 实战项目：GW Event Catalog / Glitch data 的探索性数据分析
 4. 实战项目：针对 GW150914 匹配滤波数据处理
 5. 基于 Python 的数据可视化理论与实践
 6. 实战项目：GWTC 论文中的 Figures
- 第三部分：机器学习基础
 1. 人工智能概述
 2. 机器学习的定义、目标和过程，及其常见类型
 3. 机器学习项目开发流程与准备
 4. 实战项目：对 LIGO 的 Glitch 数据实现聚类分析
- 第四部分：深度学习基础
 1. 深度学习技术概述
 2. 神经网络基础 ANN
 3. 卷积神经网络 CNN
 4. 项目实战：使用 CNN 识别双黑洞系统引力波信号
 5. 引力波数据分析与人工智能前沿

基础运维技术

- Linux 文件系统简介，权限和进程管理；创建和管理用户，设置文件权限，基本的进程管理等
- 合并冲突解决，分支管理和协作工作流程；在团队中协作一个小项目，学习如何解决合并冲突
- 基本的网络安全和 SSH 密钥管理；设置 SSH 密钥和远程访问服务器

基础容器化技术

- Docker 简介和基本操作指令；使用 Docker 创建容器应用和开发环境
- 利用 Docker 在远端服务器上，实现 Python/Jupyter 运行环境的搭建和维护
- 在 Docker 上配置合适的 Linux 编译环境，并打包包含有 LALsuite / LISAcodes 和发布软件

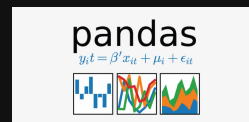




本实战训练营的课程大纲

- 第零部分：打鸡血！
- 第一部分：编程开发环境与工作流
 1. Linux 中常见指令与 shell 文本编辑
 2. Git 版本控制 (GitHub / GitLab)
 3. SSH 远程服务器 (Shell / VSCode)
 4. 基于 Docker 的容器化技术与开发环境
 5. 实战项目：Python / Jupyter 开发环境搭建
 6. 实战项目：LALsuite / LISAcode 的源码编译
- 第二部分：基于 Python 的数据分析基础
 1. Python 程序语言入门
 2. 基于 Numpy / Pandas / Scipy 的算法基础
 3. 实战项目：GW Event Catalog / Glitch data 的探索性数据分析
 4. 实战项目：针对 GW150914 匹配滤波数据处理
 5. 基于 Python 的数据可视化理论与实践
 6. 实战项目：GWTC 论文中的 Figures
- 第三部分：机器学习基础
 1. 人工智能概述
 2. 机器学习的定义、目标和过程，及其常见类型
 3. 机器学习项目开发流程与准备
 4. 实战项目：对 LIGO 的 Glitch 数据实现聚类分析
- 第四部分：深度学习基础
 1. 深度学习技术概述
 2. 神经网络基础ANN
 3. 卷积神经网络 CNN
 4. 项目实战：使用 CNN 识别双黑洞系统引力波信号
 5. 引力波数据分析与人工智能前沿

• Python 程序语言



- 常见的数据结构和高级特性；面向对象程序设计；异常处理和调试技巧
- Numpy 和 Pandas 的数据结构和常用方法；Scipy 简介
- 数据清洗与特征工程基础；线性代数等编程实现

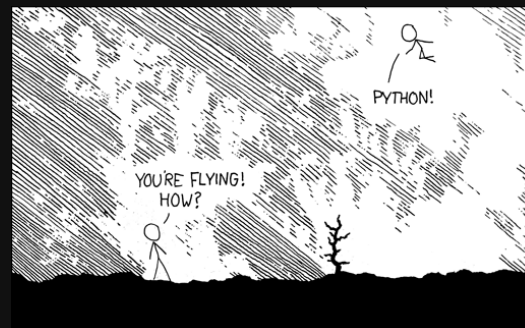
• 引力波数据处理

- 引力波数据和数字信号处理的基本操作流程



• Python 数据可视化

- 数据可视化的基本理论与美学基础；
- 基于 Matplotlib / Seaborn 的绘图编程技术
- 分析LVK合作组论文中的绘图技巧和绘图流水线的实现方法



<https://xkcd.com/353/>

matplotlib



seaborn



本实战训练营的课程大纲

- 第零部分：打鸡血！
- 第一部分：编程开发环境与 workflow
 1. Linux 中常见指令与 shell 文本编辑
 2. Git 版本控制 (GitHub / GitLab)
 3. SSH 远程服务器 (Shell / VSCode)
 4. 基于 Docker 的容器化技术与开发环境
 5. 实战项目：Python / Jupyter 开发环境搭建
 6. 实战项目：LALsuite / LISAcodes 的源码编译
- 第二部分：基于 Python 的数据分析基础
 1. Python 程序语言入门
 2. 基于 Numpy / Pandas / Scipy 的算法基础
 3. 实战项目：GW Event Catalog / Glitch data 的探索性数据分析
 4. 实战项目：针对 GW150914 匹配滤波数据处理
 5. 基于 Python 的数据可视化理论与实践
 6. 实战项目：GWTC 论文中的 Figures
- 第三部分：机器学习基础
 1. 人工智能概述
 2. 机器学习的定义、目标和过程，及其常见类型
 3. 机器学习项目开发流程与准备
 4. 实战项目：对 LIGO 的 Glitch 数据实现聚类分析
- 第四部分：深度学习基础
 1. 深度学习技术概述
 2. 神经网络基础 ANN
 3. 卷积神经网络 CNN
 4. 项目实战：使用 CNN 识别双黑洞系统引力波信号
 5. 引力波数据分析与人工智能前沿

• 人工智能

- 了解人工智能的发展背景和相关概念
- 人工智能 vs 机器学习 vs 深度学习

• 机器学习

- 机器学习的定义、目标和过程，及其常见类型
- 认识到引力波数据分析与机器学习之间的联系
- 掌握完整机器学习项目的基本流程和思维套路
- 基于 Scikit-learn 实现机器学习算法的搭建

• 深度学习

- 了解深度学习技术的发展背景和相关概念
- 人工神经网络的编程实现方法
- 卷积的定义和编程实现方法，及其与匹配滤波之间的关系
- 论文复现 CNN 识别引力波信号
- 熟悉数据预处理、模型搭建和模型验证的全流程
- 了解深度生成模型 (GAN, Flow 等) 的原理和实现方法，以及相关的引力波数据应用案例
- Kaggle 黑客马拉松

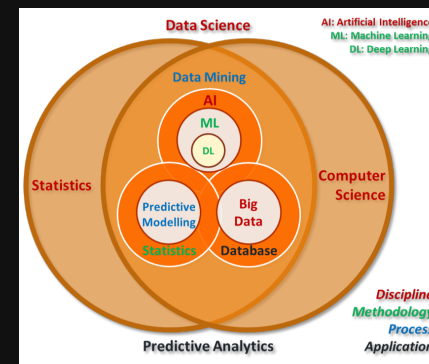




本实战训练营是什么，不是什么

• Yes

- 无需精通编程，了解最基本的本地电脑、远端服务器或集群环境的运维常识
- Python 程序设计的入门级启蒙
- Python 数据可视化的美学启蒙和学术绘图入门
- Python 数据挖掘基础
- 完整机器学习项目的基本流程和思维套路；
- 深度学习技术导论
- 深刻理解“学习”是什么意思，及其基于数据驱动的思维方式
- 基于 PyTorch 的入门级项目代码；
- 引力波数据处理与人工智能技术结合的常见科研套路
- 前沿深度学习模型的训练逻辑和引力波数据应用案例



• No

- 不是运维工程师的考证培训
- 不是算法工程师的求职培训
- 不会系统概述数据清理、特征工程等数据准备工作；
- 不是完整覆盖所有 API 等控件的 PyTorch / ... 使用指南；
- 不涉及过于细节的机器学习理论和艰深的算法原理，忽略绝大多数无关紧要的数学细节





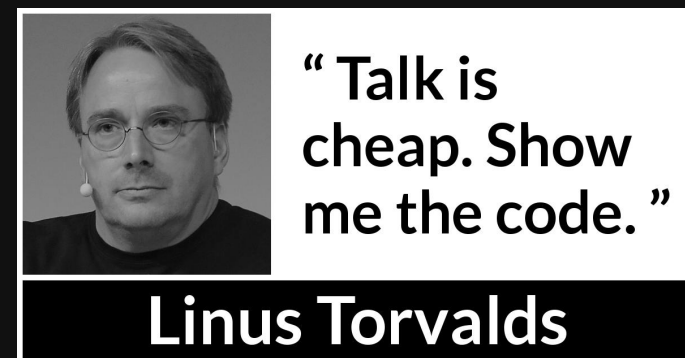
本实战训练营的学习方法与教学团队

• 如何跟随本课程学习

- 本课程的 slides / codes 等所有静态资料都会分享出来，录屏回放将限期开放
- 课程开始前，最好提前对Python编程有一些了解为最好。
 - 学会在实践中培养自己能发现问题，自己能解决问题的能力。
- 授课过程中，只全神贯注的听讲即可，不需要边听讲，边做笔记或者模仿操作
 - 目的是培养**直觉**和**联觉**；
 - 线下同学可直接提问；线上同学可在腾讯会议聊天框中留言提问；
 - 为了保证最佳听课效果，授课形式以面向**线上学员为主**；
- 课程结束后，一定要学会做技术笔记，积累个人技术财富
 - 用自己的话总结（Feynman 方法，写技术博客，...）；写给未来的自己看；
- 要坚持每天**敲代码**！
 - “...everything became much clearer when I started **writing code**.”
- **要学会自学，学会利用网络资源解决自己遇到的问题**

• 课程安排

- 灵活的授课时间：一般是在**周三、五、日的某天晚上19点**，每次上课时间将会提前一周在微信群内通知，预计2024年1月结课。
- 班主任：关于课程咨询，请找张乐老师 或 刘颖慧老师
- 助教：孙慧，赖景祺，梁博，郭意扬

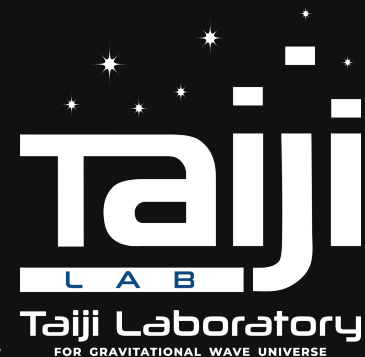
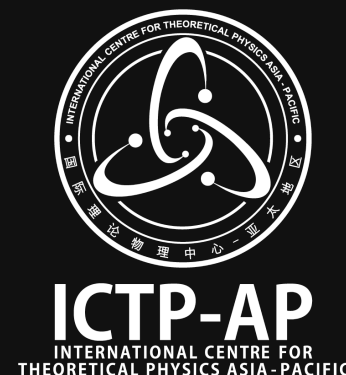




本实战训练营的考核规则和项目作业

• 考核规则

- 本课程是免费培训，**没有学分**，课程结束后的考核达标学员将会有**结业证书**发放。
- 达标要求：30%反馈问卷 + 20%考勤 + 50%基础作业
 - 反馈问卷：关于课程难度、课程内容和课程期待等
 - 考勤方法：**上线考勤**；若能按时完整完成相应的作业，也记作考勤达标
- 彩蛋：
 - 课程结束后，将会以一部分比例给优秀的达标学员赠送**小礼品**



计算服务



• 项目作业

- 本课程会有一些**【基础作业】**考验学员听课时能否快速吸收和认真听讲，难度偏低
- 另外，**【拓展作业】**将会作为与本课程内容密切相关，但需要学员自己拓展相关知识或快速自学一些技术才能完成的作业/项目。
 - 对于引力波数据处理研究方向的学员，**【扩展作业】**为必做内容。
 - 课上的实践内容与作业内容的占比大概是各占一半。
 - **【基础作业】**与**【拓展作业】**的占比也大概是各占一半。



通向自我实现之路

对“卷”的理性认识

- 人多，竞争激烈？
- 奋斗、努力、上进？
- 恶性的规则与制度。



Hi I am Guido, I created Python programming language which is used by millions of people and companies.

Interviewer: That's nice. Can we see your university degree?

Jerome de Leon @edxplere · Nov 14
Replying to @gvanrossum
I'm wondering if they still asked you to submit a resume/cv

2

2

33

Guido van Rossum @gvanrossum · Nov 14
Yes they did, and I got interviewed by Kevin Scott and Anders Hejlsberg (and others). How cool is that. :-)) (They also asked for my diploma from University?!)

14

40

183

Show replies

8:46 am · 16 Nov 20 · [Twitter Web App](#)

24 Retweets and comments 100 Likes

Matt Savino @_MattSavino · 1d
Replying to @DudeWhoCode
Did they make him implement a bubble sort on the whiteboard?

1

2



通向自我实现之路：用心卷动每一步

对“卷”的理性认识

- 人多，竞争激烈？
 - 奋斗、努力、上进？
 - 恶性的规则与制度。
-
- 心之所向，行之力卷！



Hi I am Guido, I created Python programming language which is used by millions of people and companies.

Interviewer: That's nice. Can we see your university degree?

Jerome de Leon @edxplre · Nov 14
Replying to @gvanrossum
I'm wondering if they still asked you to submit a resume/cv

2

2

33

Guido van Rossum @gvanrossum · Nov 14
Yes they did, and I got interviewed by Kevin Scott and Anders Hejlsberg (and others). How cool is that. :-) (They also asked for my diploma from University?!)

14

40

183

Show replies

8:46 am · 16 Nov 20 · [Twitter Web App](#)

24 Retweets and comments 100 Likes

Matt Savino @_MattSavino · 1d
Replying to @DudeWhoCode
Did they make him implement a bubble sort on the whiteboard?

1

2



通向自我实现之路：Evolving into a Self-Taught Expert

- 对“卷”的理性认识
- 做研究所需要的范式转换
 - 读本科 vs 读研究生：
 - 把让我做的都做了，不就行了？
 - 我该不该买本书系统性学习一下 XX？
 - 研究本身之外的必备技能/心态：
 - 英语
 - 科学上网
 - 表达
 - 速记
 - 知识管理
 - 打磨
 - 时间管理
 - ...



Public speaking
& presentations



手頭上的工作

● ● ● ● ● ● ● ●

可利用的時間

□ □ □ □ □ □ □ □

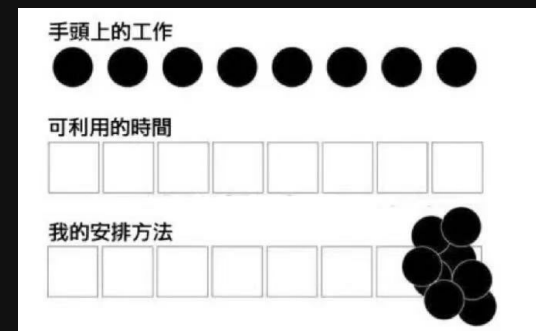
我的安排方法

□ □ □ □ □ □ □ □

● ● ● ● ● ● ● ●

通向自我实现之路：Evolving into a Self-Taught Expert

- 对“卷”的理性认识
- 做研究所需要的范式转换
 - 读本科 vs 读研究生：
 - 把让我做的都做了，不就行了？
 - 我该不该买本书系统性学习一下 XX？
 - 研究本身之外的必备技能/心态：
 - 英语
 - 科学上网
 - 表达 “输出”
 - 速记
 - 知识管理
 - 打磨
 - 时间管理
 - ...





通向自我实现之路：Teaching is Learning

- 对“卷”的理性认识
- 做研究所需要的范式转换
- 做好笔记是自我学习的不二捷径

- 笔记软件推荐: Typora, ...
- 好笔记是要用心做出来的
- 你做的笔记是写给未来的自己看的

```
16 out.write(word + " : " + str(word_count_dic[word]) + "\n") #
   写入文件
17 out.close()
18 # 文件使用完毕后必须关闭, 因为文件对象会占用操作系统的资源, 并且操作系统同一时间能打
   开的文件数量也是有限的
19
20 in_file = 'data/ShangHai.txt'
21 out_file = 'data/word_count.txt'
22 count_word_num(in_file, out_file)
```

- 由于文件读写时都有可能产生 `IOError`, 一旦出错, 后面的 `f.close()` 就不会调用。所以, Python 引入了 `with` 语句来自动帮我们调用 `close()` 方法。
- 调用 `read()` 会一次性读取文件的全部内容, 如果文件有 10G, 内存就爆了, 所以, 要保险起见, 可以反复调用 `read(size)` 方法, 每次最多读取 `size` 个字节的内容。另外, 调用 `readline()` 可以每次读取一行内容, 调用 `readlines()` 一次读取所有内容并按行返回 `list`。因此, 要根据需要决定怎么调用。

- 如果文件很小, `read()` 一次性读取最方便; 如果不能确定文件大小, 反复调用 `read(size)` 比较保险; 如果是配置文件, 调用 `readlines()` 最方便。

```
1 with open(in_file, 'r') as f: # r 表示读出
2     f.readline()
3     f.readlines()
4     for line in f.readlines():
5         print(line.strip()) # 把末尾的 '\n' 删掉
```

- 写文件时, 你可以反复调用 `write()` 来写入文件, 但是务必要调用 `f.close()` 来关闭文件。当我们写文件时, 操作系统往往不会立刻把数据写入磁盘, 而是放到内存缓存起来, 空闲的时候再慢慢写入。只有调用 `close()` 方法时, 操作系统才保证把没有写入的数据全部写入磁盘。忘记调用 `close()` 的后果是数据可能只写了一部分到磁盘, 剩下的丢失了。所以, 还是用 `with` 语句来得保险:

```
1 with open(in_file, 'w') as f:
2     f.write('Hello, world!')
```

REF: <http://www.cnblogs.com/yymjqsx/p/6554817.html>

```
1 colnames = ['date'] + [item for pair in zip(["{:02d}".format(i) for i in range(24)],
   ['flag']*24) for item in pair]
2 # 这一步在做的是搞出一个49维的数组, 其中, 第一维是日期, 第二维往后, 两个为一组, 一个是时间一个是Flag。
```

上述代码的详细解释:

- `"{:02d}".format(i)` 代表左边用 0 补位的两位数的整数来格式化为字符串。`{:}` 是 `format` 函数下格式限定符的语法, 末尾的 `2d` 表示字符串长度限定为 2, 类型为整数, `0` 为补充其余空位 (只能是一个字符, 不指定的话默认是用空格填充)。写成 `"{:0>2d}".format(i)` 效果会完全一样的, 其中 `>` 表示右对齐, 后面跟宽度。例子: `"{:0>4d}".format(55)` 表示字符串 `0055`。

See more: [Python中用format函数格式化字符串的用法](#)

- `zip()` 函数接收任意多个 (包括 0 个和 1 个) 序列列表作为参数, 返回一个 tuple 作为元素的 list 列表。关于 `zip()` 的基本运作方式, 举例胜过一切: `x = [1, 2, 3], y = [4, 5, 6], xy = zip(x, y)`, 返回的是 `[(1, 4), (2, 5), (3, 6)]`。

See more: [Python的zip函数](#)

所以, `zip = zip(["{:02d}".format(i) for i in range(24)], ['flag']*24)` 返回是 `[('00', 'flag'), ('01', 'flag'), ('02', 'flag'), ...]`

- 递推式构造列表(List Comprehension): `[item for pair in zip for item in pair]`, 其中 `zip` 是上述 `zip` 函数返回的列表。`item` 仍然是代表最终返回的列表元素, 后面的 `for...in...for...in...` 结构其实是嵌套的两个 `for` 循环:

```
1 for pair in zip:
2     for item in pair:
3         return item
```

再举一个更简单的例子, 会更加清楚:

```
1 colours = ["red", "green"]; things = ["house", "car", "tree"]
2 [(x,y) for x in colours for y in things]
3 # 返回的list是
4 # [('red', 'house'), ('red', 'car'), ('red', 'tree'), ('green', 'house'), ('green',
   'car'), ('green', 'tree')]
```

See more: [Python的递推式构造列表 \(List comprehension\)](#)



通向自我实现之路：学会搜索答案

- 对“卷”的理性认识
- 做研究所需要的范式转换
- 做好笔记是自我学习的不二捷径
- 自学路上必会之“上下求索”的技能

- 正确的态度：不存在能够知道一切答案并能轻松提供给你的人，但所有与答案相关的信息都一定能在互联网上找到蛛丝马迹。
- 要充分利用百度和Google等搜索引擎
- 基本搜索技巧：

“关键词取的越多，显示的网页信息会越准确相关，但是真正相关的网页条目越少。关键词取的越少，自然地搜索结果会越泛。”

```
hello world | hello world | hello world |
File "C:\frozen importlib\_bootstrap", line 646, in _load_unlocked
File "C:\frozen importlib\_bootstrap", line 616, in _load_backward_compatible
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2985, in <module>
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2971, in _call_aside
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 3013, in _initialize_master_wor
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2544, in activate
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2118, in declare_namespace
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2058, in _handle_ns
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2087, in _rebuild_mod_path
AttributeError: '_NamespacePath' object has no attribute 'sort'

Process finished with exit code 1
```

pandas 随机行

网页 新闻 贴吧 知道 音乐 图片 视频 地图 文库 更多»

百度为您找到相关结果约415,000个

python-Pandas学习 如何对数据集随机抽样? - SCUT_Sam - CSDN博客

2017年5月2日 - 摘要:有时候我们只需要数据集中的一部分,并不需要全部的数据。这个时候我们就要对数据集进行随机的抽样。pandas中自带有抽样的方法。应用场景:我有10W...
blog.csdn.net/qq_22238... - 百度快照

Python-Pandas 如何shuffle(打乱)数据? - SCUT_Sam - CSDN博客

2017年4月28日 - 其中参数frac是要返回的比例,比如df中有10行数据,我只想返回其中的30%,那么...上一篇朴素贝叶斯入门 下一篇python-Pandas学习 如何对数据集随机抽样?...
blog.csdn.net/qq_22238... - 百度快照

pandas - python如何随机且有放回地从训练集中的抽取N个训练样本...

2016年8月25日 - python如何随机且有放回地从训练集中的抽取N个训练样本 pandas numpy
python...训练集行数为5000 2016年08月25日提问 评论 邀请回答 编辑 举报 提交...
https://segmentfault.com/q/101... - 百度快照

pandas随机选取若干行 - baidu_15113429的博客 - CSDN博客

2017年10月16日 - 在实验交叉验证的时候,可以在pandas中使用sample进行随机选取。import
pandas as pd dic = {'a':[1, 2, 3, 4], 'b':[5, 6, 7, 8], 'c':[9,...
blog.csdn.net/baidu_15... - 百度快照

attributeerror: '_namespacepath' object has no attribute 'sort'

全部 视频 新闻 图片 更多 设置 工具

找到约 589 条结果 (用时 0.83 秒)

小提示: 仅限搜索简体中文结果。您可以在设置中指定搜索语言

[_NamespacePath object has no attribute sort \(31.0.0\) · Issue #885 ...](#)
<https://github.com/pyppa/pip/issues/885> · 翻译此页
2016年12月12日 - File "c:\akitruntime\lib\site-packages\pip_vendor\pkg_resources__init__.py", line 2121, in _rebuild_mod_path orig_path.sort(key=position_in_sys_path) AttributeError: '_NamespacePath' object has no attribute 'sort'. Setuptools is listed as version 34.3.3 inside the virtualenv lib\site-packages. I hope this ...

[NamespacePath object has no attribute sort #4216 - GitHub](#)
<https://github.com/pyppa/pip/issues/4216> · 翻译此页
2017年1月5日 - I am relieving this : AttributeError: '_NamespacePath' object has no attribute 'sort' I read everything you wrote at this discussion and others with similar problems and users. I tried all suggested codes, but always receiving errors or syntax error or command not found. I am working on Mac, but tried on ...

[AttributeError: '_NamespacePath' object has no attribute 'sort' · Issue ...](#)
<https://github.com/GoogleCloudPlatform/google-cloud-python/issues/2990> · 翻译此页
2017年2月8日 - Ubuntu 16.04.1 LTS; Python 3.5.2 :: Anaconda 4.2.0 (64-bit). I was trying install with pip install --upgrade google-cloud but I get the following AttributeError: '_NamespacePath' object has no attribute 'sort'. Below is the stack trace: Traceback (most recent call last): File
"home/ubuntu/anaconda3/bin/pip", line ...

[_NamespacePath object has no attribute sort - conda · Issue #5798 ...](#)
<https://github.com/conda/conda/issues/5798> · 翻译此页
2017年8月9日 - _rebuild_mod_path(path, packageName, module) File "C:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources__init__.py", line 2087, in _rebuild_mod_path orig_path.sort(key=position_in_sys_path) AttributeError: '_NamespacePath' object has no attribute 'sort' ...

[python - AttributeError: '_NamespacePath' object has no attribute ...](#)
[https://stackoverflow.com/questions/46111111/attributeerror-namespacepath-object-has-no-attr...](https://stackoverflow.com/questions/46111111/attributeerror-namespacepath-object-has-no-attr) · 翻译此页



通向自我实现之路：学会搜索答案

- 对“卷”的理性认识
- 做研究所需要的范式转换
- 做好笔记是自我学习的不二捷径
- 自学路上必会之“上下求索”的技能

- 正确的态度：不存在能够知道一切答案并能轻松提供给你的人，但所有与答案相关的信息都一定能在互联网上找到蛛丝马迹。
- 要充分利用百度和Google等搜索引擎
- 基本搜索技巧：

“关键词取的越多，显示的网页信息会越准确相关，但是真正相关的网页条目越少。关键词取的越少，自然地搜索结果会越泛。”

“只要它存在，就没有查不到的！
因为善于搜索，所以了解世界。

Because it's there!

```
hello world | hello world | hello world
File "C:\frozen\importlib\_bootstrap.py", line 646, in _load_unlocked
File "C:\frozen\importlib\_bootstrap.py", line 616, in _load_backward_compatible
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2985, in <module>
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2971, in _call_aside
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 3013, in _initialize_master_wor
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2544, in activate
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2118, in declare_namespace
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2058, in _handle_ns
File "D:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources\__init__.py", line 2087, in _rebuild_mod_path
AttributeError: '_NamespacePath' object has no attribute 'sort'

Process finished with exit code 1
```

pandas 随机行

网页 新闻 贴吧 知道 音乐 图片 视频 地图 文库 更多»

百度为您找到相关结果约415,000个

[python-Pandas学习 如何对数据集随机抽样? - SCUT_Sam - CSDN博客](#)
2017年5月2日 - 摘要:有时候我们只需要数据集中的一部分,并不需要全部的数据。这个时候我们就要对数据集进行随机的抽样。pandas中自带有的抽样的方法。应用场景:我有10W...
[blog.csdn.net/qz_22238...](#) - 百度快照

[Python-Pandas 如何shuffle\(打乱\)数据? - SCUT_Sam - CSDN博客](#)
2017年4月28日 - 其中参数frac是要返回的比例,比如df中有10行数据,我只想返回其中的30%,那么...上一篇朴素贝叶斯入门 下一篇python-Pandas学习 如何对数据集随机抽样?...
[blog.csdn.net/qz_22238...](#) - 百度快照

[pandas - python如何随机且有放回地从训练集中的抽取N个训练样本...](#)
2016年8月25日 - python如何随机且有放回地从训练集中的抽取N个训练样本 pandas numpy
python...训练集行数为5000 2016年08月25日提问 评论 邀请回答 编辑 举报 提交...
[https://segmentfault.com/q/101...](#) - 百度快照

[pandas随机选取若干行 - baidu_15113429的博文 - CSDN博客](#)
2017年10月16日 - 在实验交叉验证的时候,可以在pandas中使用sample进行随机选取。import
pandas as pd dic = {'a':[1, 2, 3, 4], 'b':[5, 6, 7, 8], 'c':[9,...
[blog.csdn.net/baidu_15...](#) - 百度快照

attributeerror: '_namespacepath' object has no attribute 'sort'

全部 视频 新闻 图片 更多 设置 工具

找到约 589 条结果 (用时 0.83 秒)

小提示: 仅限搜索简体中文结果。您可以在设置中指定搜索语言

[_NamespacePath object has no attribute sort \(31.0.0\) · Issue #885 ...](#)
[https://github.com/pyppa/setuputils/issues/885](#) - 翻译此页
2016年12月12日 - File "C:\akitruntime\lib\site-packages\pip_vendor\pkg_resources__init__.py", line 2121, in _rebuild_mod_path orig_path.sort(key=position_in_sys_path) AttributeError: '_NamespacePath' object has no attribute 'sort'. Setuputils is listed as version 34.3.3 inside the virtualenv lib/site-packages. I hope this ...

[NamespacePath object has no attribute sort #4216 - GitHub](#)
[https://github.com/pyppa/pip/issues/4216](#) - 翻译此页
2017年1月5日 - I am relieving this: AttributeError: '_NamespacePath' object has no attribute 'sort' I read everything you wrote at this discussion and others with similar problems and users. I tried all suggested codes, but always receiving errors or syntax error or command not found. I am working on Mac, but tried on ...

[AttributeError: '_NamespacePath' object has no attribute 'sort' · Issue ...](#)
[https://github.com/GoogleCloudPlatform/google-cloud-python/issues/2990](#) - 翻译此页
2017年2月8日 - Ubuntu 16.04.1 LTS; Python 3.5.2 :: Anaconda 4.2.0 (64-bit). I was trying install with pip install --upgrade google-cloud but I get the following AttributeError: '_NamespacePath' object has no attribute 'sort'. Below is the stack trace: Traceback (most recent call last): File "/home/ubuntu/anaconda3/bin/pip", line ...

[_NamespacePath object has no attribute sort - conda · Issue #5798 ...](#)
[https://github.com/conda/conda/issues/5798](#) - 翻译此页
2017年2月8日 - _rebuild_mod_path(path, packageName, module) File "C:\ProgramData\Anaconda3\lib\site-packages\setuptools-27.2.0-py3.6.egg\pkg_resources__init__.py", line 2087, in _rebuild_mod_path orig_path.sort(key=position_in_sys_path) AttributeError: '_NamespacePath' object has no attribute 'sort' ...

[python - AttributeError: '_NamespacePath' object has no attribute ...](#)
[https://stackoverflow.com/.../attributeerror-namespacepath-object-has-no-attrib...](#) - 翻译此页



通向自我实现之路：学会请教问题

- 对“卷”的理性认识
- 做研究所需要的范式转换
- 做好笔记是自我学习的不二捷径
- 自学路上必会之“上下求索”的技能
- 自学路上必会之“提问的艺术”

问一下，群里有人会用caffe的python api生成prototxt文件嘛

生成网络模型

最近在这方面遇到一点问题

1月9日 15:59

群里的高手可否交流一下



20:38:42

请问有研究CNN的吗哈

不写paper
不改头像

20:39:17

直接丢问题

不写paper
不改头像

20:39:23

这么问不会有回应的

09:35:17
大神，xgboost按照说明安装好，python3环境也检验成功，但是import就是不行啊，请问是什么原因啊？

09:35:25

```
ModuleNotFoundError                               Traceback (most recent call last)
<ipython-input-1-2f3326c353e> in <module>()
    18
    19 #导入xgboost
--> 20 import xgboost as xgb
    21 from xgboost import XGBClassifier
    22

ModuleNotFoundError: No module named 'xgboost'
```



通向自我实现之路：学会请教问题

- 对“卷”的理性认识
- 做研究所需要的范式转换
- 做好笔记是自我学习的不二捷径
- 自学路上必会之“上下求索”的技能
- 自学路上必会之“提问的艺术”
 - 要知道，只有真正清楚和了解自己是遇到了什么问题，才能给出合理和清晰的提问。
 - 当你不知道自己遇到什么问题的时候，先了解一下该如何提问。
 - Fast.ai 的《如何提问的技巧》
 - 另外有一个 informative 版本的《提问的智慧》

“提问的内容尽量包含：

1. 你为了解决这个问题，你是做了哪些操作？你希望这些操作是什么效果和实际的效果是什么？
2. 你猜测这个问题的根源是什么，以及你认为解决该问题可能需要如何做？
3. 请告知你的基本电脑环境：如什么操作系统？在云平台？你的库版本是啥等～
4. 如果你的操作步骤不同于课程给的notebook模板，请强调出来～
5. 你收到的报错信息是什么（很重要！）
6. 如果可以，请**截图**！
7. 如果和课程中操作的返回结果不一致，请指明～
8. 你已经解决的部分问题是怎么做的？比方说你重启了电脑？重启了kernel？
9. 你发觉到某些可行的解决方案，可以贴出来，如来自csdn、简书等～

提问环节

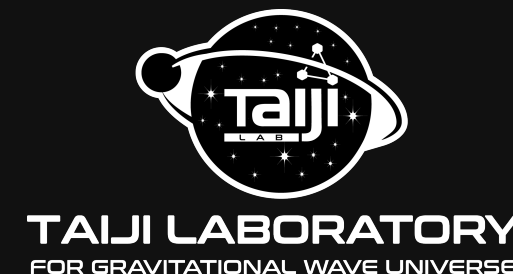


中国科学院大学
University of Chinese Academy of Sciences

- 请关注公众号，了解更多信息
 - 国际理论物理中心（亚太地区）



- 主办单位
 - 中国科学院大学·国际理论物理中心（亚太地区）
 - 引力波宇宙太极实验室



- 赞助单位
 - 中科曙光

