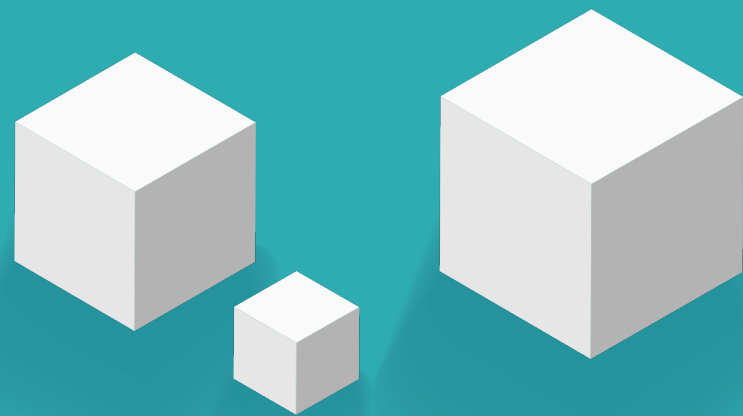
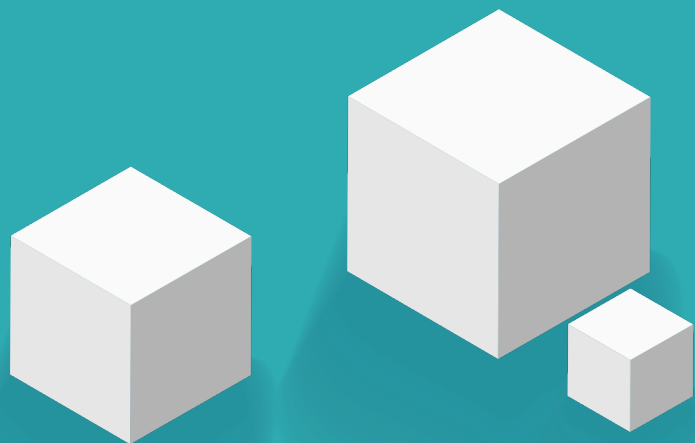


《机器学习在天文中的应用》

Application of machine learning in astronomy

郝 奇

南京大学天文与空间科学学院



《机器学习在天文中的应用》课程介绍

Contents

1

课程目标

2

教学大纲

3

课程内容

4

教学方式

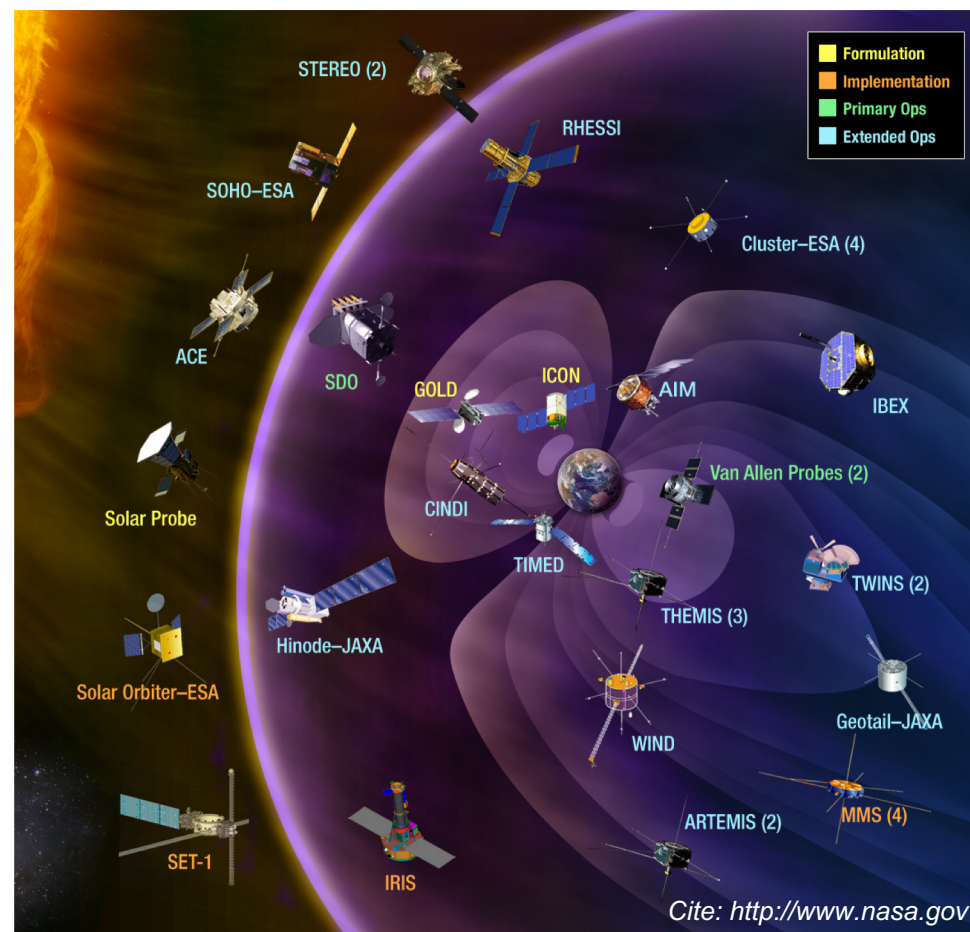
/01

课程目标

- 课程设计目的
- 课程目标

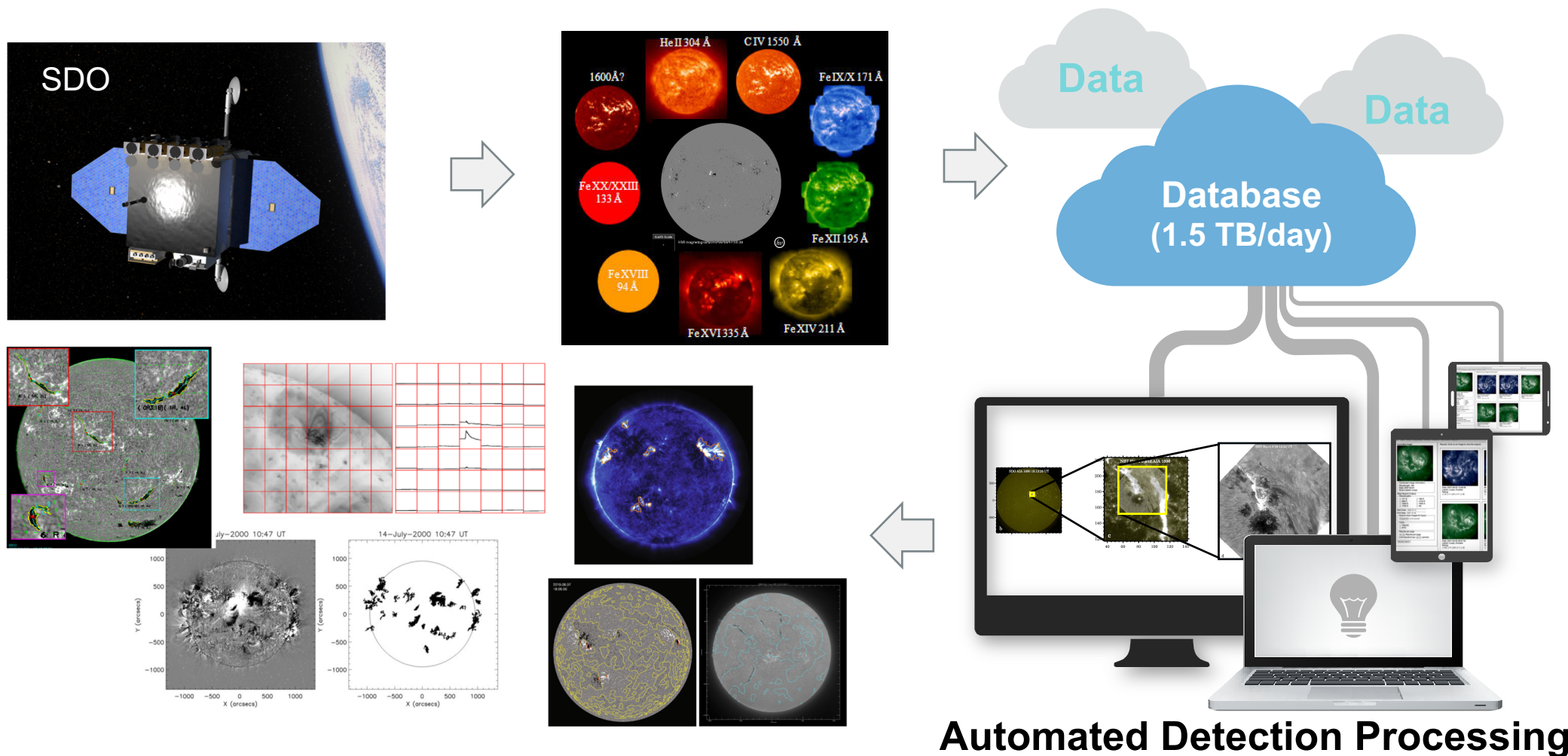
0. 课程背景

High resolution observations result in vast data



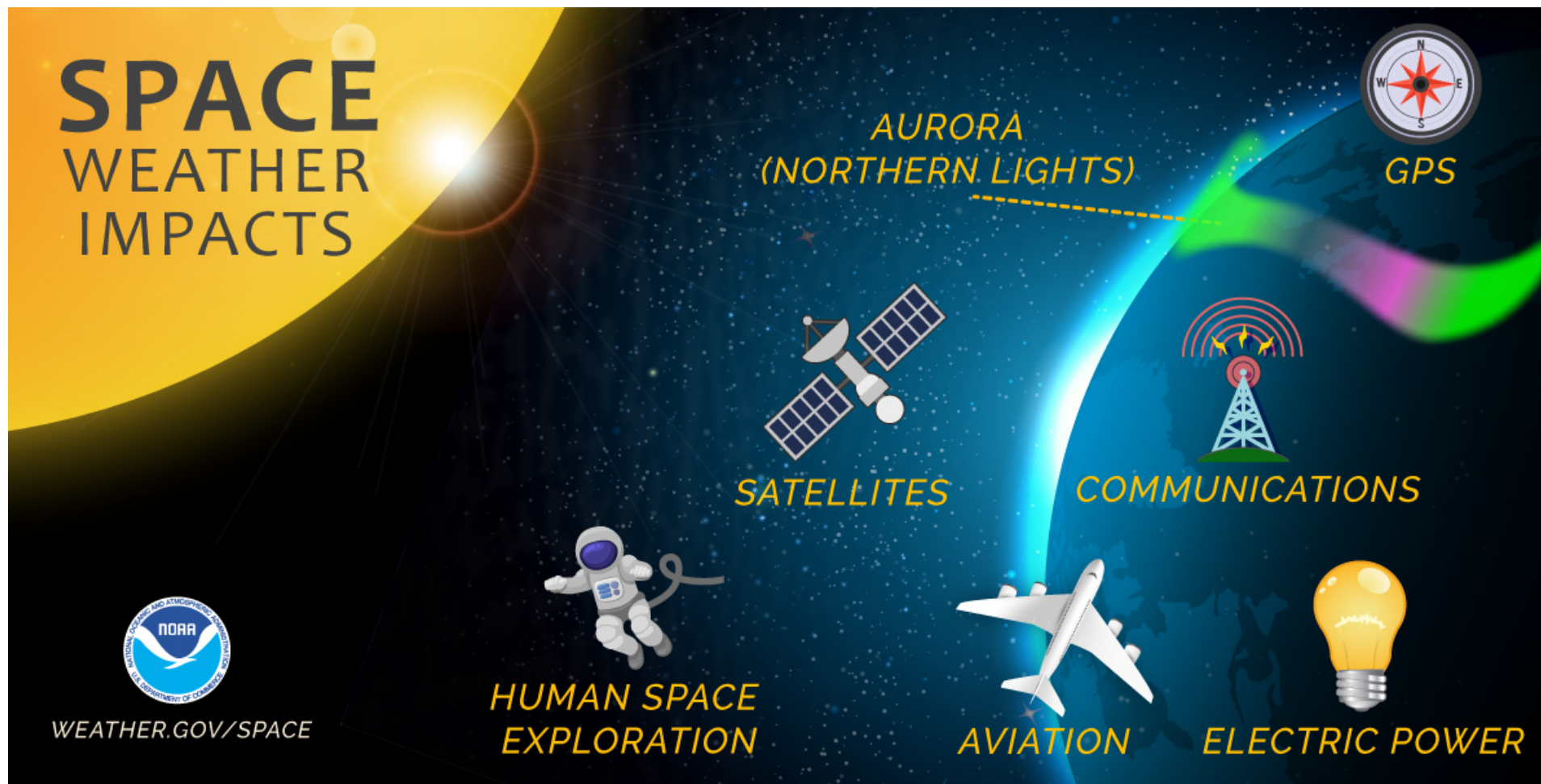
0. 课程背景

Fast derive features of interest



0. 课程背景

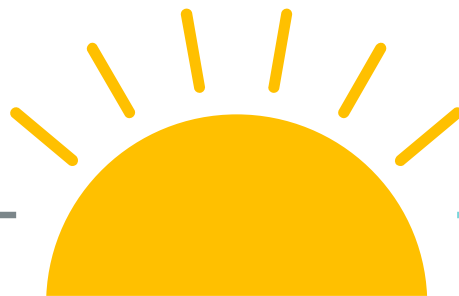
Forecast the space weather



1. 课程目标

天文大数据时代

天文进入大数据时代，海量观测数据处理需要专业的处理技术



课程与科研衔接

本科生进行早期科研训练、研究生进行课题研究，无法从课程学习状态直接进入课题研究状态，对数据一筹莫展

为什么要开设这门课？

专业需求

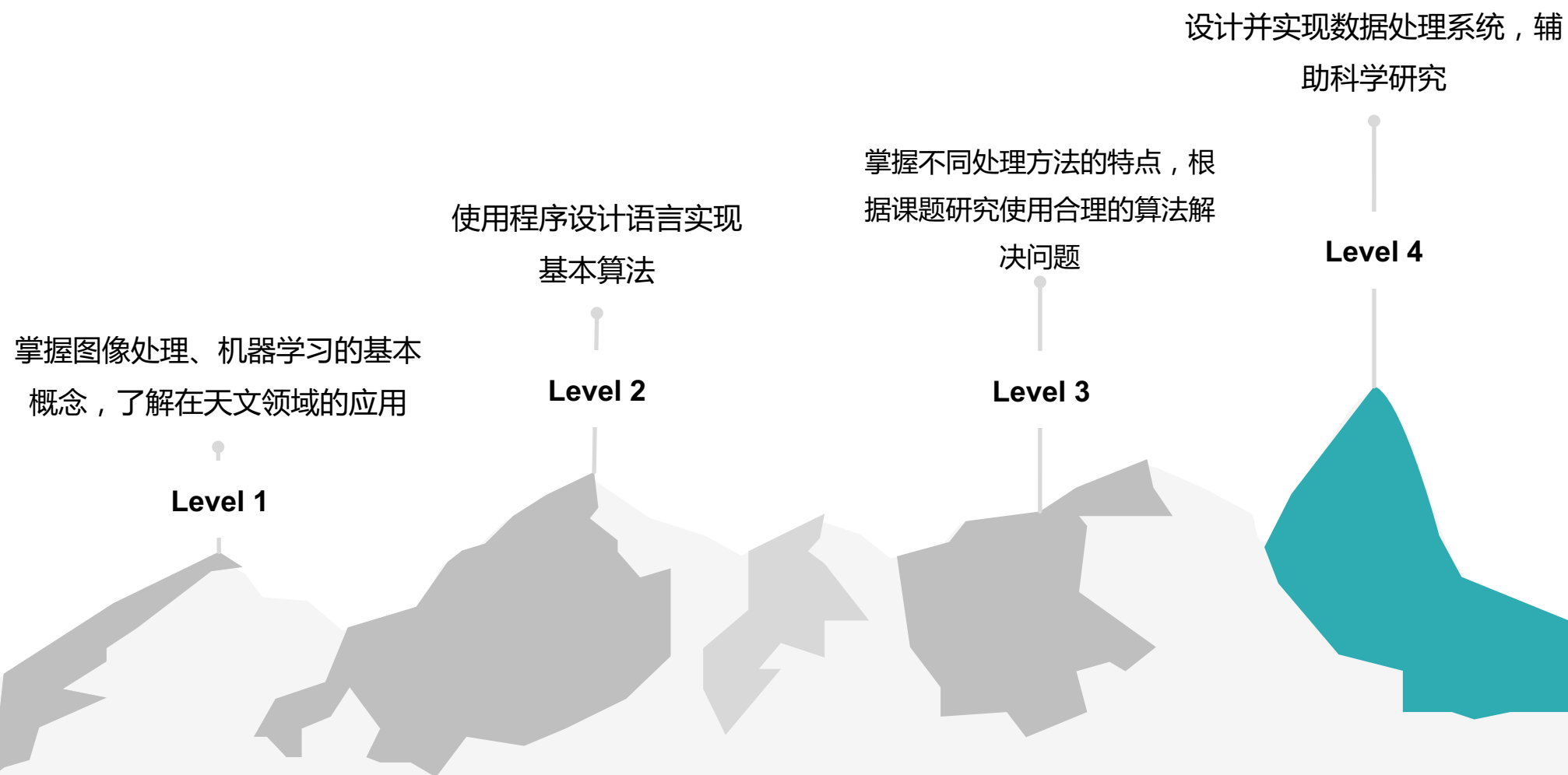
天文学科学生未接受系统的图像处理、机器学习方法的知识学习，无法快速准确地针对自己的研究课题找到合适的数据处理方法



研究趋势

从早期信号处理到图像处理，再到图像识别、机器学习等各种人工智能方法开始广泛应用于天文领域

1. 课程目标

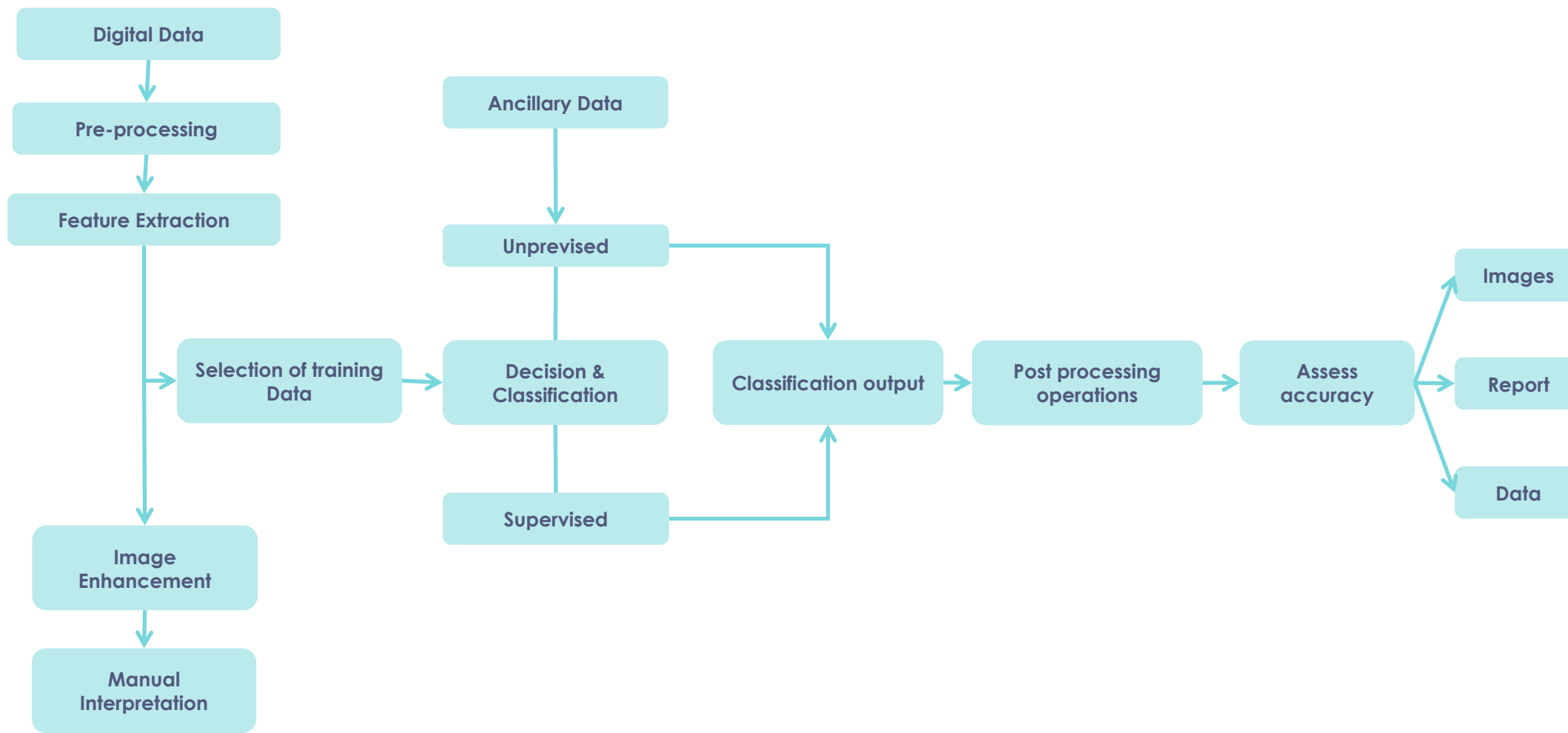


/02

教学大纲

- 课程整体概况
- 课程大纲

2. 教学大纲



2. 教学大纲

- 图像处理（Image processing）指用计算机对图像进行分析，以达到所需结果的技术。它是计算机科学的一门专业课，主要讲授图像处理的数字处理技术，包括处理、变换与分析的基本理论、方法和技术。
- 机器学习（Machine learning）是一种从数据中自动分析获得规律，并利用规律对未知数据进行处理的过程。它是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、计算复杂性理论等多门学科。
- 机器学习在天文中的应用，则以天文数据为中心，侧重图像处理技术、机器学习方法在天文数据处理中的应用。

第一部分

- 数字图像的基本概念
- 数字图像基本处理方法

第二部分

- 图像增强、图像分割与图像识别等在天文领域的具体应用



第三部分

- 机器学习的基本概念
- 回归、SVM、集成学习、神经网络模型等介绍

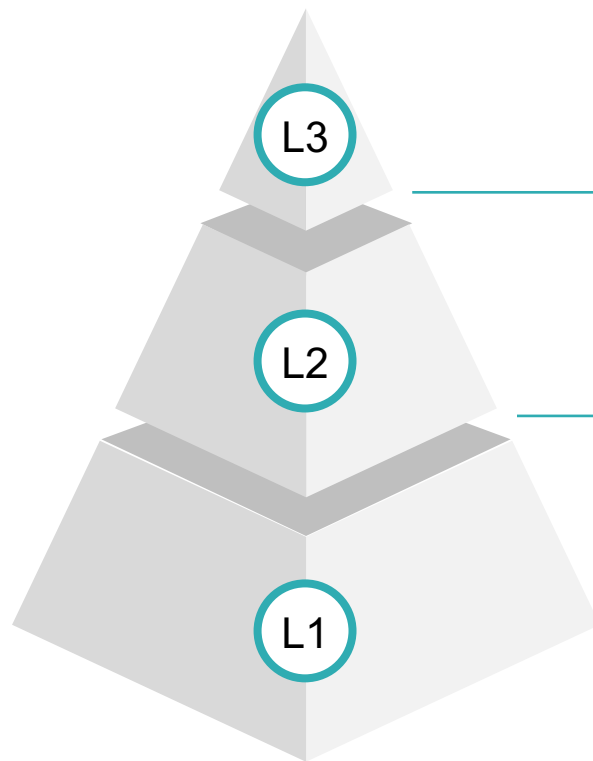
第四部分

- 机器学习在天文中的应用
（主要以太阳物理为例）

2. 教学大纲

授课对象：本科三、四年级；研究生一年级

已掌握一定的基础知识，能够使用编程语言进行程序设计及编写，即将或正在从事天文课题研究。



- 《人工智能》
- 《计算机图形学》
- 《数据挖掘》

- 《模式识别》
- 《计算机视觉》

• 《机器学习在天文中的应用》

- 《高等数学》（必修）
- 《线性代数》（必修）
- 《程序设计》（必修）

- 《概率统计》（必修）
- 《数据结构》（可选）
- 《离散数学》（可选）

/03

课程内容

3. 课程内容

第一部分（基本概念）

- ☐ 图像处理概述
- ☐ 灰度直方图
- ☐ 点运算
- ☐ 代数运算
- ☐ 几何运算
- ☐ 邻域运算
- ☐ 形态学运算

第二部分（天文应用）

- ☐ 线性系统及其应用
- ☐ 傅立叶/小波变换及其应用
- ☐ 图像增强、对齐、复原应用
- ☐ 滤波器设计及应用
- ☐ 基于图像处理的自动探测

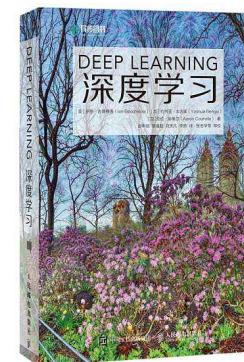
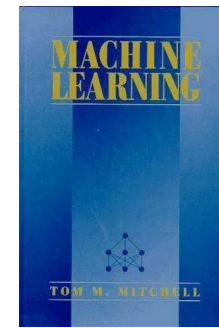
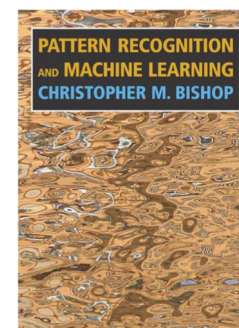
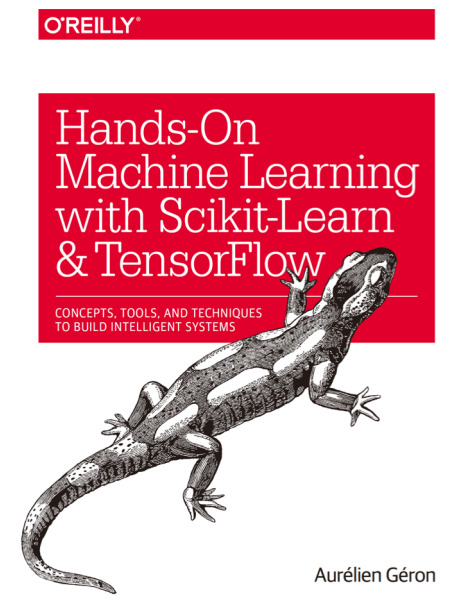
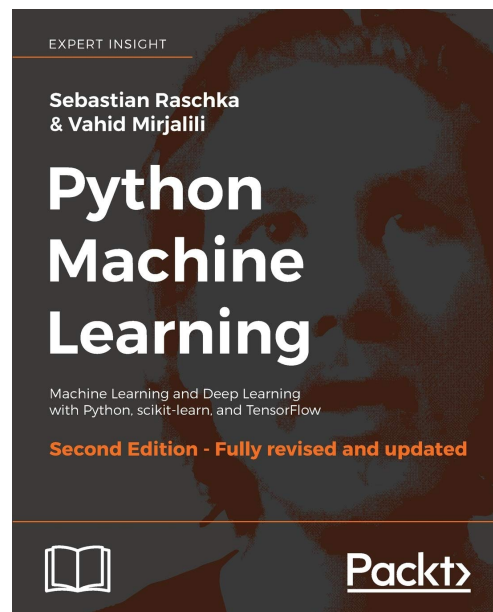
第三部分（算法介绍）

- ☐ 回归
- ☐ 支持向量机（SVM）
- ☐ 决策树
- ☐ 概率生成模型
- ☐ 神经网络
- ☐ 集成学习
- ☐ 聚类、降维与度量学习

第四部分（天文应用）

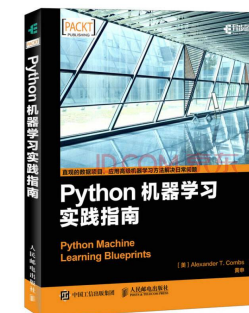
- ☐ 线性回归在天文中的应用
- ☐ SVM分类在天文中的应用
- ☐ 决策树在天文中的应用
- ☐ 概率生成模型在天文中的应用
- ☐ 神经网络在天文中的应用
- ☐ 其他机器学习方法在天文中的应用

3. 课程内容



1. 《机器学习》，周志华，清华大学出版社；
2. Python机器学习（原书第2版），Sebastian Raschka；
3. Hands-On Machine Learning with Scikit-Learn and TensorFlow, Aurélien Géron；
4. 《统计学习方法（第2版）》，李航，清华大学出版社；

.....



/04

教学方式

4. 教学方式

01 教师

- 课堂讲授（幻灯片，图像、视频、音频素材）
- 实验指导（上机操作，课堂/课后？）

02 学生

- 课堂学习，掌握基本概念（基本要求）
- 课后思考，学会使用不同算法处理不同数据（目标）

03 实验

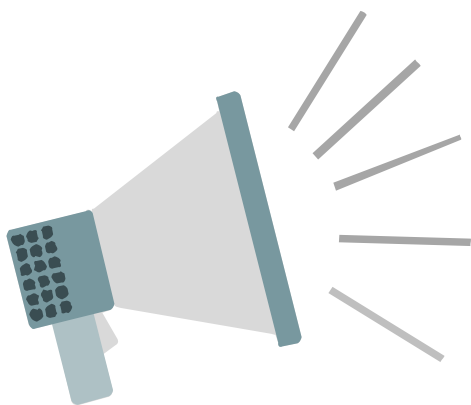
- 针对基础概念，尝试使用擅长的编程语言实现基本算法（3个左右？）
- 根据自己科研课题（早期科研、研究课题项目等），设计课题

04 讨论

- 课堂讨论
- 小组讨论，针对课题设计讨论等



4. 教学方式



学生主导

课程只有2/3的时间是教师讲授，其他时间由学生主导，针对预设课程实验或者学生设计的课题，进行研讨学习



动手动脑

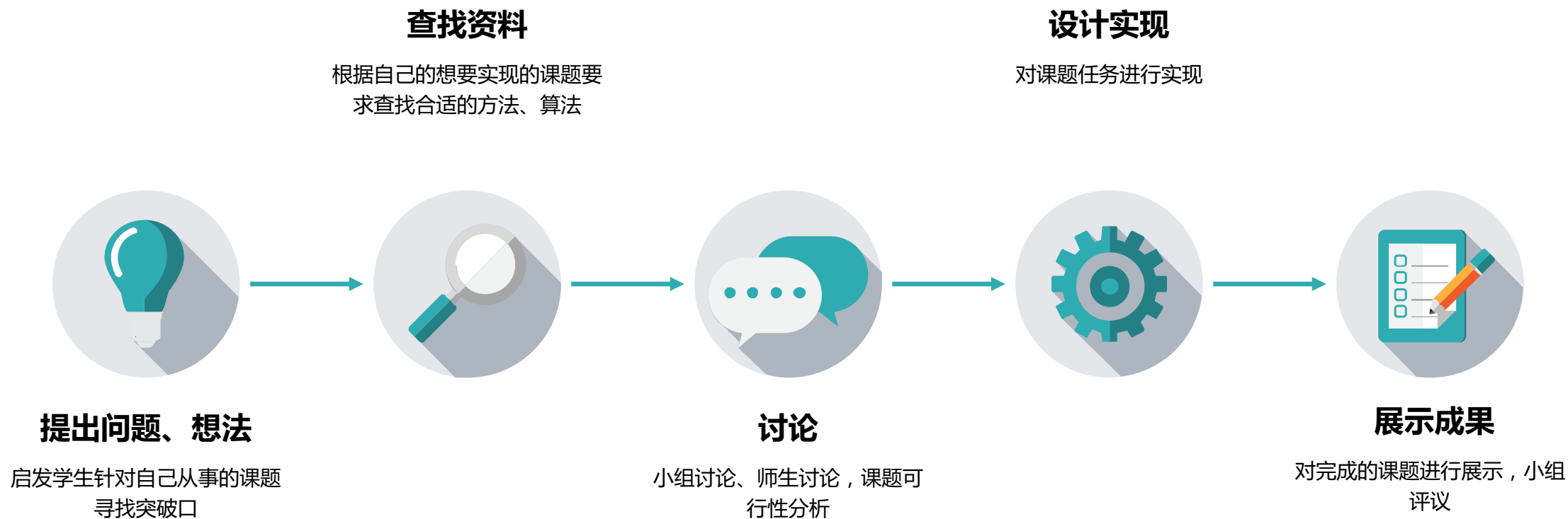
课程不是单纯的讲与听，需要课堂或课后进行讨论、实验、编写程序、查找资料及文献



小组合作

课程设置小组模式，鼓励学生合作，通过讨论、研讨完成课程实验、报告以及课题

4. 教学方式



4. 教学方式

平时（随堂测验）20%，作业（习题报告）30%，期末报告（论文）50%



4. 教学方式

教 师：郝 奇

办公室：天文楼 431室

电 话：89681932（办公室）

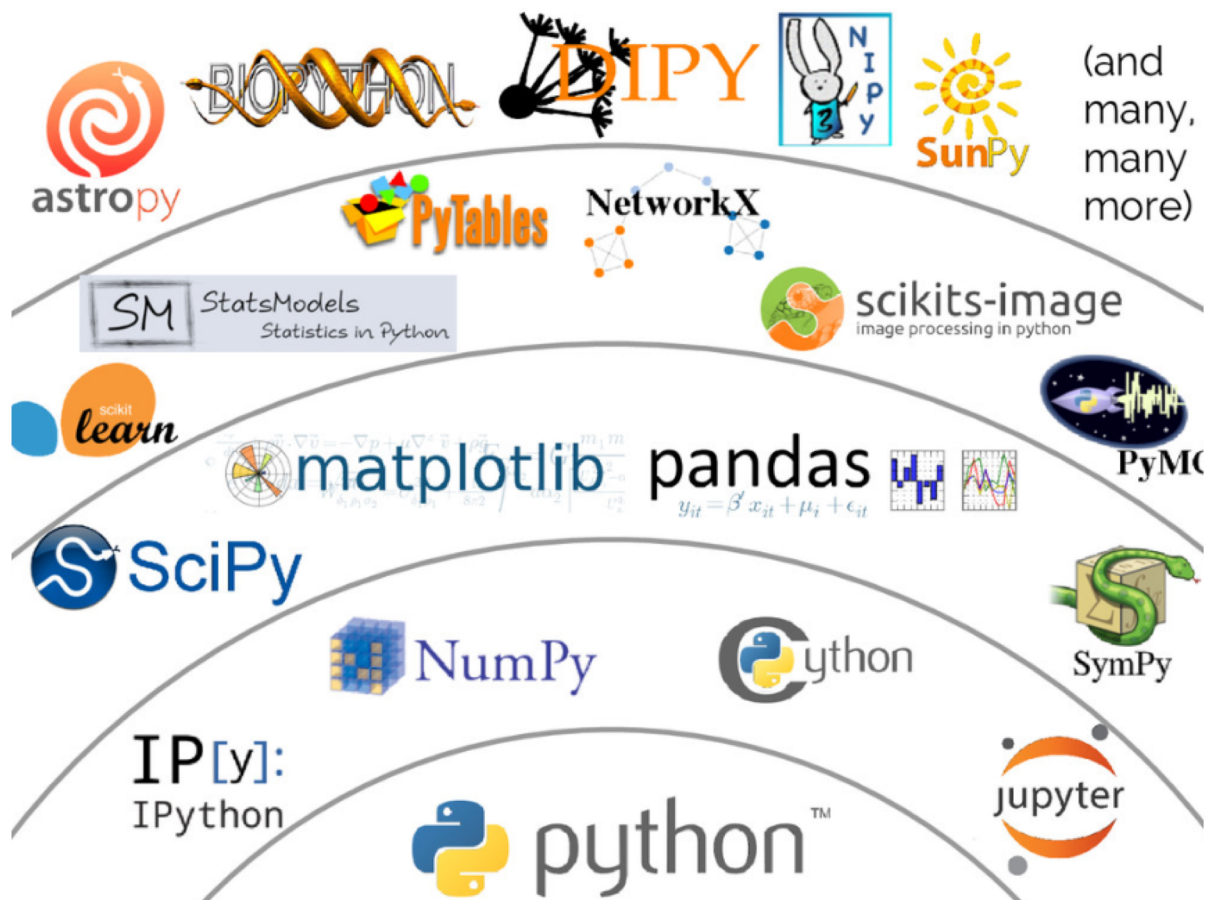
邮 箱：haoqi@nju.edu.cn

课后讨论时间：TBA.

助 教：



4. 教学方式



TIOBE Index for September 2020

Sep 2020	Sep 2019	Change	Programming Language	Ratings	Change
1	2	▲	C	15.95%	+0.74%
2	1	▼	Java	13.48%	-3.18%
3	3		Python	10.47%	+0.59%
Sep 2020	Sep 2019	Change	Programming Language		
4					
5					
6	1	2	▲	C	
7					
8	2	1	▼	Java	
9	3	3		Python	
10					
11	4	4		C++	
12					
13	5	5		C#	
14	6	6		Visual Basic	
15					
16	7	7		JavaScript	
17					
18	8	9	▲	PHP	
19	9	19	▲	R	
20					
	10	8	▼	SQL	

4. 教学方式



The Jupyter Notebook

JupyterLab

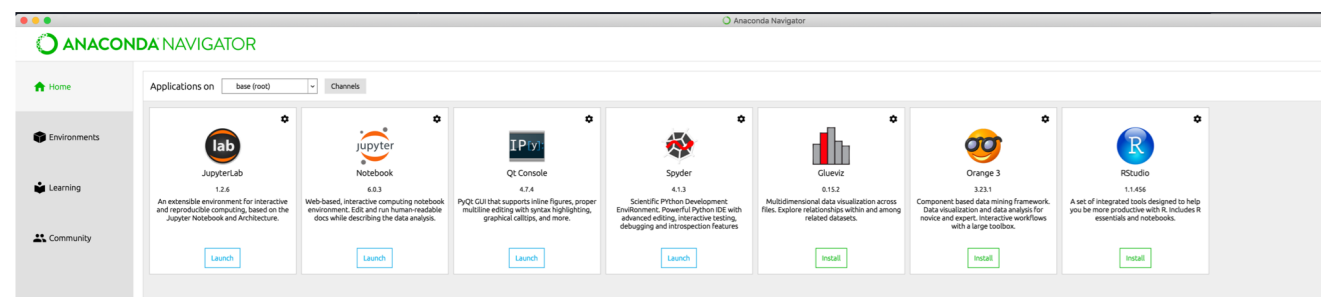
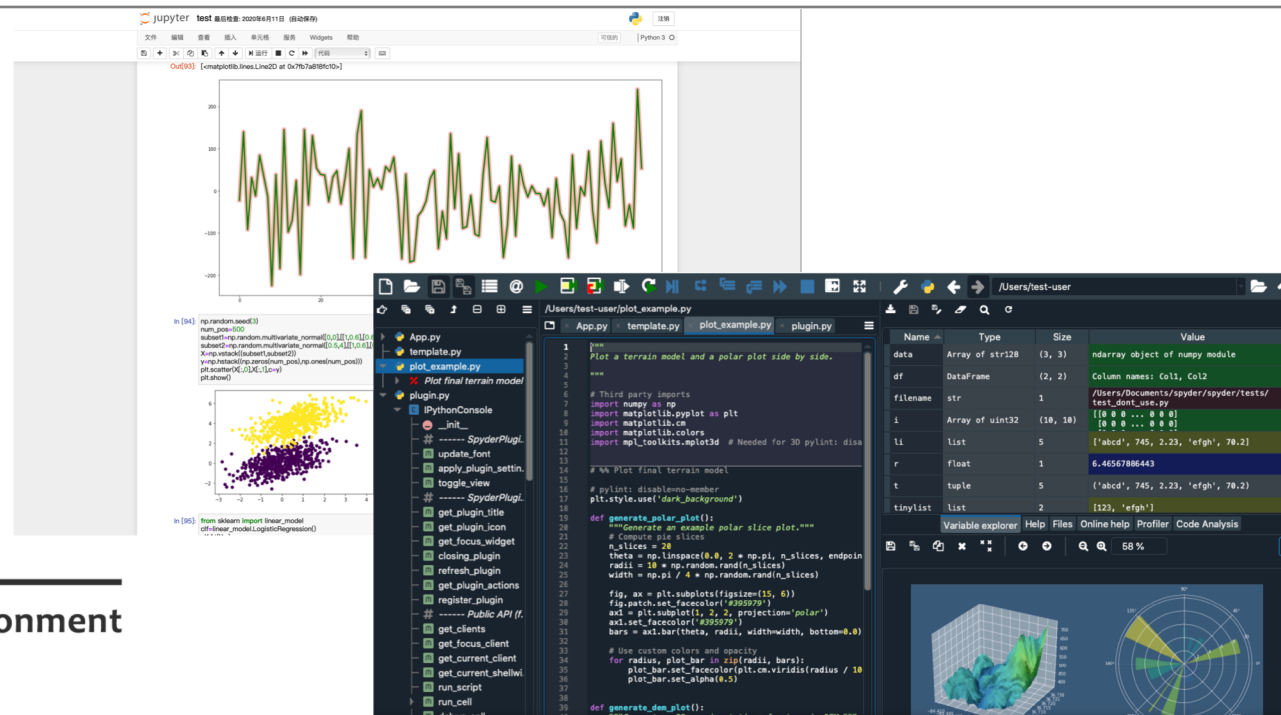


SPYDER

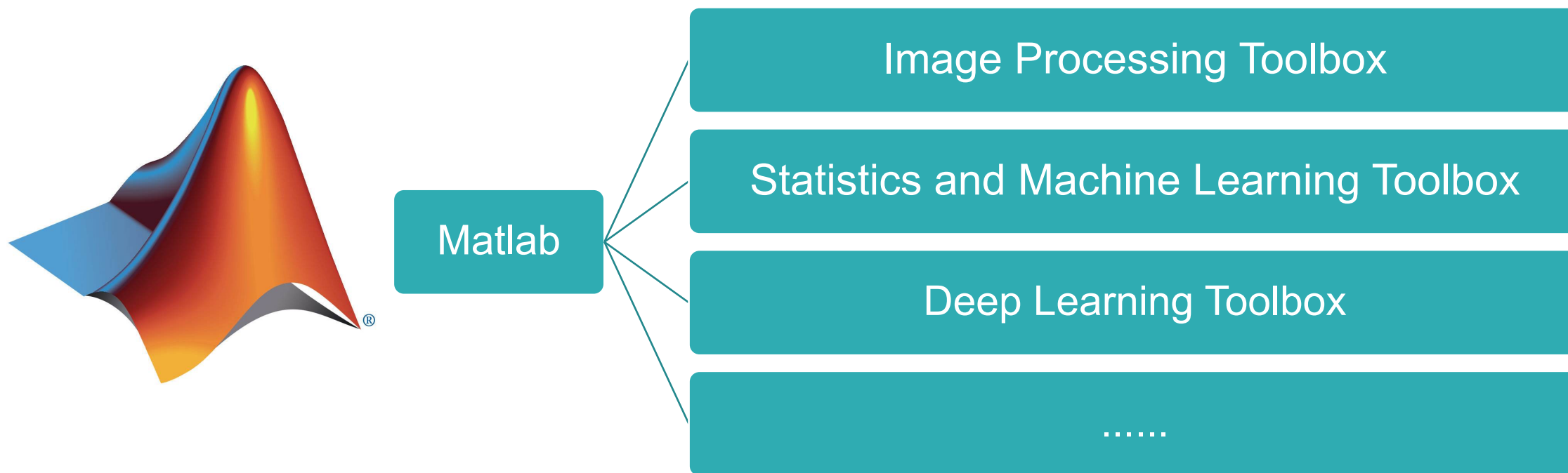
The Scientific Python Development Environment

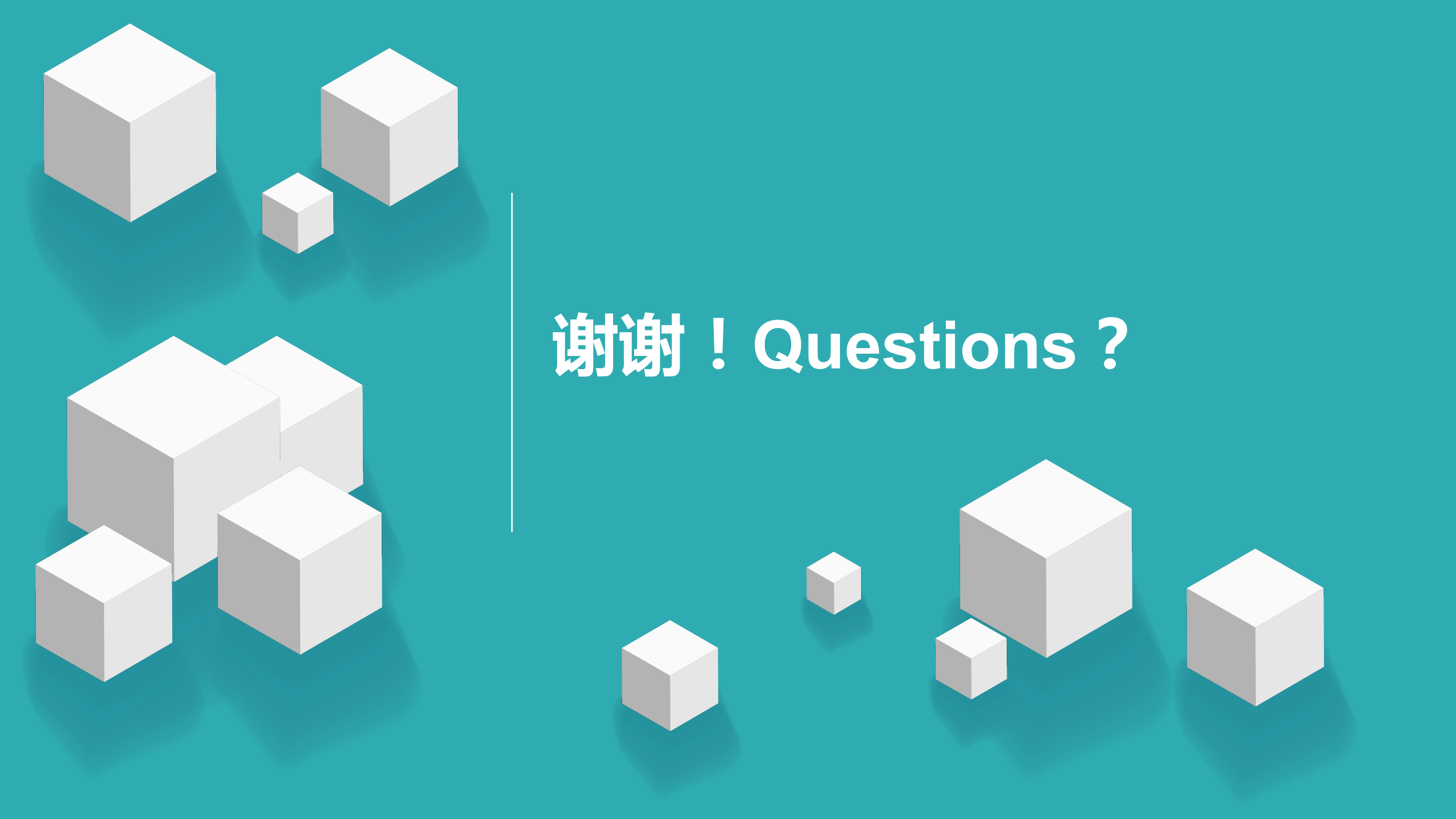


ANACONDA CLOUD



4. 教学方式



The background is a solid teal color. It is decorated with several white 3D cubes of various sizes. In the top left, there are three cubes: two large ones and one small one in between. In the bottom left, there is a cluster of four cubes, including two large ones and two smaller ones. In the bottom center, there is a single medium-sized cube. To the right of the center, there is a small cube. Further right, there is a group consisting of a large cube, a small cube next to it, and another medium-sized cube to the right. On the far right, there is a single medium-sized cube. A thin white vertical line is positioned to the left of the central text.

谢谢！Questions？