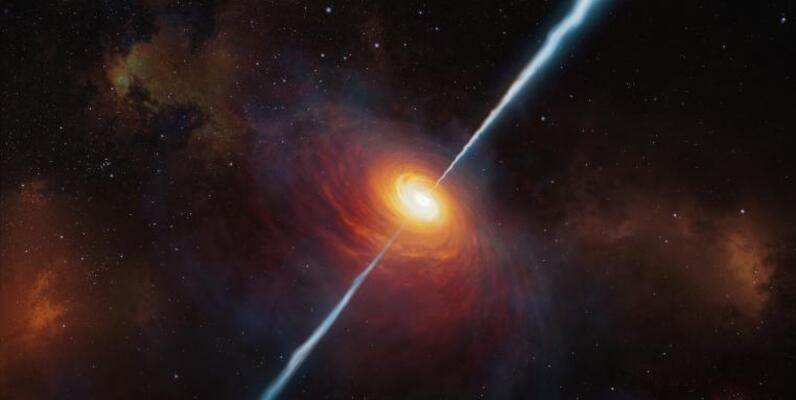
**深空天体分类**

恒星、星系和类星体的分类。



在天文学中，恒星分类是根据恒星的光谱特征对其进行分类。星系、类星体和恒星的分类方案是天文学中最基本的分类方案之一。早期对恒星及其在天空中的分布进行的编目使人们了解到，它们构成了我们自己的星系。在仙女座与我们自己的银河系分离之后，随着更强大的望远镜的建造，许多星系开始被观测。该数据集旨在根据恒星、星系和类星体的光谱特征对其进行分类。

这些数据由对宇宙深空目标进行的100000次观测组成。每一次观测都由17个特征列和1个分类列描述，这些分类列将其确定为恒星、星系或类星体。

本项目的主要目标是建立一个机器学习模型， 分类列将其确定为恒星、星系或类星体。

1. 理解数据，清洗数据
2. 随机选取一部分数据作为训练集，一部分作为验证集，一部分作为测试集
3. 设计机器学习分类算法并编程实现
4. 攥写完整的报告

obj\_ID=对象标识符，CAS使用的图像目录中标识对象的唯一值

α=右旋上升角（J2000历元）

delta=倾角（J2000历元）

u=光度系统中的紫外线滤光片

g=光度系统中的绿色滤光片

r=光度系统中的红色滤光片

i=光度系统中的近红外滤光片

z=光度系统中的红外滤光片

run\_ID=用于识别特定扫描的运行编号

Rerun\_ID=重新运行编号以指定图像的处理方式

cam\_col=摄像机列，用于识别运行中的扫描线

field\_ID=用于识别每个字段的字段号

spec\_obj\_ID=用于光学光谱对象的唯一ID（这意味着具有相同spec\_obj-ID的两个不同观测输出同样分类）

class=物体类别（星系、恒星或类星体物体）

红移=基于波长增加的红移值

plate= plate ID，观测中的每个plate编号

MJD=修改后的儒略日期，用于指示何时获取给定的观测数据。

儒略日（Julian Day）是在儒略周期内以连续的日数计算时间的计时法，主要是天文学家在使用。儒略日数（Julian Day Number，JDN）的计算是从格林威治标准时间的中午开始，包含一个整天的时间，起点的时间（0日）回溯至儒略历的公元前4713年1月1日中午12点（在格里历是公元前4714年11月24日），这个日期是三种多年周期的共同起点，且是历史上最接近现代的一个起点。例如，2000年1月1日的UT12:00是儒略日2,451,545。

fiber\_ID=光纤ID，用于标识在每次观察中将光指向焦平面的光纤